

Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on: July 29, 2004, 10:21:25 ; Search time: 2490 Seconds

(without all alignments)
1235,263 Million cell updates/sec

29: gb_gss2:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

	%	Query Score	Match Length	DB ID	Description
Result No.					
1	478	86.1	594	12	BM292030 EST574562
2	475	85.6	718	12	BM290603 EST577137
3	472	85.0	491	12	BG226380 kg2ch09_Y
4	469	84.5	590	12	BM291076 EST577610
5	453	81.6	715	12	BM291800 EST574342
6	440.5	79.4	593	13	BX673604 BX673604
7	440.5	79.4	600	12	BM251446 BOTL01000
C	8	440.5	79.4	628	10 BE881329 501490517
9	440.5	79.4	998	11	BC019532 Mus muscu
10	440.5	79.4	2530	11	AK047413 Mus muscu
11	427.5	77.0	448	9	AV616467 AV616467
12	427.5	77.0	459	10	AW204560 UI-H-B11-
13	427.5	77.0	492	12	BI181558 UNL-P-FN-
14	427.5	77.0	505	9	AI804396 tcs62909_X
15	427.5	77.0	545	9	AI382221 tdd4c05_X
16	427.5	77.0	553	9	AI698758 wa1d11_X
17	427.5	77.0	556	10	AW298104 UI-H-BW0-
18	427.5	77.0	593	9	AI766667 w1d2e06_X
19	427.5	77.0	681	10	BF057343 7K19e01_X
20	427.5	77.0	716	13	BQ603828 ML-P-CP1-
21	427.5	77.0	722	9	AI439635 tcs91d08_X
22	427.5	77.0	741	10	BP058976 tcs36f08_X
23	427.5	77.0	753	9	AI808294 wf54b07_X
24	427.5	77.0	822	9	AI418347 tcs48b07_X
25	427.5	77.0	848	9	AI522901 tcs991910_X
26	427.5	77.0	870	12	BI182367 UNL-P-FN-
27	427.5	77.0	883	13	BO604124 ML-P-CP1-
28	427.5	77.0	894	9	AI472322 tcs17c05_X
29	427.5	77.0	896	12	BI183684 UNL-P-FN-
30	425.5	76.7	580	9	AI761654 wh99a04_X
31	425.5	76.7	743	13	BX297137 BX297137
32	423.5	76.3	648	10	BF115229 tcs80b07_X
33	422.5	76.1	613	9	AI221705 qn05a04_X
34	421.5	75.9	608	10	AW573095 tcs13h12_X
35	421.5	75.9	612	10	BF057618 tcs46h02_X
36	421.5	75.9	621	10	AW517242 tcs92g08_X
37	421.5	75.9	686	10	BF057595 tcs46e06_X
38	421.5	75.9	757	9	AI813640 w1ts5a09_X
39	421.5	75.9	793	12	BI183274 UNL-P-FN-
40	421.5	75.9	813	12	BI184900 UNL-P-FN-
41	419.5	75.9	397	14	CB77245 AMGNNUC:T
42	419.5	75.6	422	9	AA769305 n239e07_S
43	419.5	75.6	685	14	CA316065 UI-M-FW0-
44	418.5	75.4	577	9	AW028783 w735b10_X
45	418.5	75.4	677	12	BI185172 UNL-P-FN-
					ALIGMENTS
RESULT 1					
LOCUS	BM292020				
DEFINITION	EST574562 AvSG Amblyomma variegatum cDNA clone AVABF26'				
SEQUENCE					
ACCESSION	BM292020				
VERSION	BM292020.1				
EST.					
KEYWORDS					
SOURCE	Amblyomma variegatum				
ORGANISM	Amblyomma variegatum				
Eukaryota; Metazoa; Arthropoda; Arachnida; Ixodida; Amblyomma.					
Parasitiformes; Ixodida; Acari;					
REFERENCE	1 (bases 1 to 594)				

29: gb_gss2:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

	%	Query Score	Match Length	DB ID	Description
Result No.					
1	478	86.1	594	12	BM292030 EST574562
2	475	85.6	718	12	BM290603 EST577137
3	472	85.0	491	12	BG226380 kg2ch09_Y
4	469	84.5	590	12	BM291076 EST577610
5	453	81.6	715	12	BM291800 EST574342
6	440.5	79.4	593	13	BX673604 BX673604
7	440.5	79.4	600	12	BM251446 BOTL01000
C	8	440.5	79.4	628	10 BE881329 501490517
9	440.5	79.4	998	11	BC019532 Mus muscu
10	440.5	79.4	2530	11	AK047413 Mus muscu
11	427.5	77.0	448	9	AV616467 AV616467
12	427.5	77.0	459	10	AW204560 UI-H-B11-
13	427.5	77.0	492	12	BI181558 UNL-P-FN-
14	427.5	77.0	505	9	AI804396 tcs62909_X
15	427.5	77.0	545	9	AI382221 tdd4c05_X
16	427.5	77.0	553	9	AI698758 wa1d11_X
17	427.5	77.0	556	10	AW298104 UI-H-BW0-
18	427.5	77.0	593	9	AI766667 w1d2e06_X
19	427.5	77.0	681	10	BF057343 7K19e01_X
20	427.5	77.0	716	13	BQ603828 ML-P-CP1-
21	427.5	77.0	722	9	AI439635 tcs91d08_X
22	427.5	77.0	741	10	BP058976 tcs36f08_X
23	427.5	77.0	753	9	AI808294 wf54b07_X
24	427.5	77.0	822	9	AI418347 tcs48b07_X
25	427.5	77.0	848	9	AI522901 tcs991910_X
26	427.5	77.0	870	12	BI182367 UNL-P-FN-
27	427.5	77.0	883	13	BO604124 ML-P-CP1-
28	427.5	77.0	894	9	AI472322 tcs17c05_X
29	427.5	77.0	896	12	BI183684 UNL-P-FN-
30	425.5	76.7	580	9	AI761654 wh99a04_X
31	425.5	76.7	743	13	BX297137 BX297137
32	423.5	76.3	648	10	BF115229 tcs80b07_X
33	422.5	76.1	613	9	AI221705 qn05a04_X
34	421.5	75.9	608	10	AW573095 tcs13h12_X
35	421.5	75.9	612	10	BF057618 tcs46h02_X
36	421.5	75.9	621	10	AW517242 tcs92g08_X
37	421.5	75.9	686	10	BF057595 tcs46e06_X
38	421.5	75.9	757	9	AI813640 w1ts5a09_X
39	421.5	75.9	793	12	BI183274 UNL-P-FN-
40	421.5	75.9	813	12	BI184900 UNL-P-FN-
41	419.5	75.9	397	14	CB77245 AMGNNUC:T
42	419.5	75.6	422	9	AA769305 n239e07_S
43	419.5	75.6	685	14	CA316065 UI-M-FW0-
44	418.5	75.4	577	9	AW028783 w735b10_X
45	418.5	75.4	677	12	BI185172 UNL-P-FN-
					ALIGMENTS
RESULT 1					
LOCUS	BM292020				
DEFINITION	EST574562 AvSG Amblyomma variegatum cDNA clone AVABF26'				
SEQUENCE					
ACCESSION	BM292020				
VERSION	BM292020.1				
EST.					
KEYWORDS					
SOURCE	Amblyomma variegatum				
ORGANISM	Amblyomma variegatum				
Eukaryota; Metazoa; Arthropoda; Arachnida; Ixodida; Amblyomma.					
Parasitiformes; Ixodida; Acari;					
REFERENCE	1 (bases 1 to 594)				

AUTHORS	Nene, V., Lee, D., Quackenbush, J., Skilton, R., Mwaura, S., Gardner, M.J. and Bishop, R.
TITLE	AvGI, an index of genes transcribed in the salivary glands of the ixodid tick Amblyomma variegatum
JOURNAL	Int. J. Parasitol. 32 (12), 1447-1456 (2002)
MEDLINE	2281296
PUBMED	12392910
COMMENT	Contact: Vish Nene Parasite Genomics Group The Institute for Genomic Research 9712 Medical Center Drive, Rockville, MD 20850, USA Tel: 301-610-5968 Fax: 301-838-0208 Email: nene@tigr.org Seq primer: M13 reverse. Location/Qualifiers 1. .594 /organism="Amblyomma variegatum" /mol_type="mRNA" /db_xref="taxon:34610" /clone="AVAB26" /tissue_type="Salivary glands" /dev_stage="Adult" /lab_host="E.coli strain DH10B-TonA" /clone_lib="AvSG"
FEATURES	/note="Vector: pCMV-SPORT6.1; Salivary glands were dissected on day five after initiation of feeding. Total RNA was prepared using acid guanidium thiocyanate-phenol-chloroform extraction. The cDNA library was custom prepared by Invitrogen Corporation. Briefly, first strand cDNA was primed using oligo(dT) containing a NotI site. Size fractionated double stranded cDNA was ligated to EcoRV-NotI cleaved vector and electroporated into E.coli."
SOURCE	source
ORIGIN	S-09-844-353A-54 (1-103) × BM292020 (1-594) Alignment Scores: 1 LysLysThrThrArgArgAsnAlaTrpGlyAsnNetSerTyraGluLeuIleThr 20 259 AAAAGAACCTGGCTGCCATTTCAGGCAAGGGCACAGCAACAGCTCCGTGAAAG 318 b b Percent Similarity: 94.06% Local Similarity: 83.17% Query Match: 86.13% BPs: 12 Length: 594 Matches: 84 Conservative: 11 Mismatches: 6 Index: 0 Gaps: 0
Y Y	1 LysLysThrThrArgArgAsnAlaTrpGlyAsnNetSerTyraGluLeuIleThr 20 379 GTGCACAACTGGCTGCCATTTCAGGCAAGGGCACAGCAACAGCTCCGTGAAAG 438 b b Percent Similarity: 94.06% Local Similarity: 83.17% Query Match: 86.13% BPs: 12 Length: 594 Matches: 84 Conservative: 11 Mismatches: 6 Index: 0 Gaps: 0
Y Y	21 ThrAlaIleMetAlaSerProGluIysArgLeuThrLeuIleGlnValTygGlutrpMet 40 319 CAAGCUCATCAAAGCGCCGAGAACGCTCACAGATCTAGTGATGATG 378 b b Percent Similarity: 94.06% Local Similarity: 83.17% Query Match: 86.13% BPs: 12 Length: 594 Matches: 84 Conservative: 11 Mismatches: 6 Index: 0 Gaps: 0
Y Y	41 ValGlnAsnValProTrpPheArgAspSerAsnSerAlaGlyTrpLys 60 379 GTGCACAACTGGCTGCCATTTCAGGCAAGGGCACAGCAACAGCTCCGTGAAAG 438 b b Percent Similarity: 94.06% Local Similarity: 83.17% Query Match: 86.13% BPs: 12 Length: 594 Matches: 84 Conservative: 11 Mismatches: 6 Index: 0 Gaps: 0
Y Y	61 AsnSerIleArgHisAsnLeuSerLeuIleHisSerArgPheMetAsnProArg 100 439 AACTCTATCCGGCACACCTGCTGCAAGCATGCGATTCAGCTGAGTGAAG 498 b b Percent Similarity: 94.06% Local Similarity: 83.17% Query Match: 86.13% BPs: 12 Length: 594 Matches: 84 Conservative: 11 Mismatches: 6 Index: 0 Gaps: 0
Y Y	81 AlaGlyLysSerSerTrpTrpValIleAsnProAspAlaSerGlnGluGly 80 499 GCGGGAAAGAGTCGGTGAATGCTGAACCCGGCAAGCCGAGGGCGAGG 558 b b Percent Similarity: 94.06% Local Similarity: 83.17% Query Match: 86.13% BPs: 12 Length: 594 Matches: 84 Conservative: 11 Mismatches: 6 Index: 0 Gaps: 0
Y Y	101 Arg 101 559 CGC 561
DEFINITION	EST577137 AvSG Amblyomma variegatum cDNA clone AVAA069 5' end, mRNA sequence.
ACCESSION	BM290603
VERSION	1
KEYWORDS	EST, Amblyomma variegatum
SOURCE	Organism: Amblyomma variegatum Eukaryota: Metazoa; Arthropoda; Chelicera; Arachnida; Acari;
ORGANISM	Parasitiformes; Ixodida; Ixodidae; Amblyomma.
REFERENCE	1 (bases 1 to 718)
AUTHORS	Nene, V., Lee, D., Quackenbush, J., Skilton, R., Mwaura, S., Gardner, M.J. and Bishop, R.
TITLE	AvGI, an index of Genes transcribed in the salivary glands of the ixodid tick Amblyomma variegatum
JOURNAL	Int. J. Parasitol. 32 (12), 1447-1456 (2002)
MEDLINE	2281296
PUBMED	12392910
COMMENT	Contact: Vish Nene Parasite Genomics Group The Institute for Genomic Research 9712 Medical Center Drive, Rockville, MD 20850, USA Tel: 301-610-5968 Fax: 301-838-0208 Email: nene@tigr.org Seq primer: M13 reverse.
LOCATION/QUALIFIERS	Location/Qualifiers 1. .594 /organism="Amblyomma variegatum" /mol_type="mRNA" /db_xref="taxon:34610" /clone="AVAA069" /tissue_type="Salivary glands" /dev_stage="Adult" /lab_host="E.coli strain DH10B-TonA" /clone_lib="AvSG" /note="Vector: pCMV-SPORT6.1; Salivary glands were dissected on day five after initiation of feeding. Total RNA was prepared using acid guanidium thiocyanate-phenol-chloroform extraction. The cDNA library was custom prepared by Invitrogen Corporation. Briefly, first strand cDNA was primed using oligo(dT) containing a NotI site. Size fractionated double stranded cDNA was ligated to EcoRV-NotI cleaved vector and electroporated into E.coli."
FEATURES	source
PRIMER	Seq Primer: M13 reverse.
LOCATION/QUALIFIERS	Location/Qualifiers 1. .718 /organism="Amblyomma variegatum" /mol_type="mRNA" /db_xref="taxon:34610" /clone="AVAA069" /tissue_type="Salivary glands" /dev_stage="Adult" /lab_host="E.coli strain DH10B-TonA" /clone_lib="AvSG" /note="Vector: pCMV-SPORT6.1; Salivary glands were dissected on day five after initiation of feeding. Total RNA was prepared using acid guanidium thiocyanate-phenol-chloroform extraction. The cDNA library was custom prepared by Invitrogen Corporation. Briefly, first strand cDNA was primed using oligo(dT) containing a NotI site. Size fractionated double stranded cDNA was ligated to EcoRV-NotI cleaved vector and electroporated into E.coli."
ORIGIN	
ALIGNMENT SCORES:	
Y Y	Score: 4.66e-47 Qy 1 LysLysThrThrArgArgAsnAlaTrpGlyAsnNetSerTyraGluLeuIleThr 20 Db 307 AAAAGAACCTGCTGCCATTTCAGGCAAGGGCACAGCAACAGCTCCGTGCAAGCCCTAGGCGAACATGCTCTAGGCCACAGTACAGTGATGACCTGATACG 366 b b Percent Similarity: 93.07% Best Local Similarity: 83.17% Query Match: 85.59% DB: 12 Length: 718 Matches: 84 Conservative: 10 Mismatches: 7 Index: 0 Gaps: 0
Y Y	21 ThrAlaIleMetAlaSerProGluIysArgLeuThrLeuIleGlnValTygGlutrpMet 40 Qy 21 ThrAlaIleMetAlaSerProGluIysArgLeuThrLeuIleGlnValTygGlutrpMet 40 Db 367 CAAGCUCATCAAAGCGCCGAGAACGCTCCAGAACGCTACAGTGATGACCTGATACG 426 b b Percent Similarity: 93.07% Best Local Similarity: 83.17% Query Match: 85.59% DB: 12 Length: 718 Matches: 84 Conservative: 10 Mismatches: 7 Index: 0 Gaps: 0
Y Y	41 ValGlnAsnValProTrpPheArgAspSerAsnSerAlaGlyTrpLys 60 Qy 41 ValGlnAsnValProTrpPheArgAspSerAsnSerAlaGlyTrpLys 60 Db 427 GTGCAGAACGCTGCCATTTCAGGCAAGGGCACAGCAACAGCTCCGTGCAAGCCCTAGGCGAACATGCTCTAGGCCACAGTACAGTGATGACCTGATACG 486 b b Percent Similarity: 93.07% Best Local Similarity: 83.17% Query Match: 85.59% DB: 12 Length: 718 Matches: 84 Conservative: 10 Mismatches: 7 Index: 0 Gaps: 0
Y Y	61 AsnSerIleArgHisAsnLeuSerLeuIleHisSerArgPheMetAsnProArg 100 Qy 61 AsnSerIleArgHisAsnLeuSerLeuIleHisSerArgPheMetAsnProArg 100 Db 487 AACTCTATCCGGCACACCTGCTGCAAGCATGCGATTCAGGCAAGCATGCTGACAGTACAGTGATGACCTGATACG 546 b b Percent Similarity: 93.07% Best Local Similarity: 83.17% Query Match: 85.59% DB: 12 Length: 718 Matches: 84 Conservative: 10 Mismatches: 7 Index: 0 Gaps: 0
Y Y	81 AlaGlyLysSerSerTrpTrpValIleAsnProAspAlaSerGlnGluGly 80 Qy 81 AlaGlyLysSerSerTrpTrpValIleAsnProAspAlaSerGlnGluGly 80 Db 81 AlaGlyLysSerSerTrpTrpValIleAsnProAspAlaSerGlnGluGly 80 b b Percent Similarity: 93.07% Best Local Similarity: 83.17% Query Match: 85.59% DB: 12 Length: 718 Matches: 84 Conservative: 10 Mismatches: 7 Index: 0 Gaps: 0
Y Y	101 Arg 101 Qy 101 Arg 101 Db 101 Arg 101 b b Percent Similarity: 93.07% Best Local Similarity: 83.17% Query Match: 85.59% DB: 12 Length: 718 Matches: 84 Conservative: 10 Mismatches: 7 Index: 0 Gaps: 0
DEFINITION	EST577137 AvSG Amblyomma variegatum cDNA clone AVAA069 5' end, mRNA sequence.
ACCESSION	BM290603
VERSION	1
KEYWORDS	EST, Amblyomma variegatum
SOURCE	Organism: Amblyomma variegatum Eukaryota: Metazoa; Arthropoda; Chelicera; Arachnida; Acari;
ORGANISM	Parasitiformes; Ixodida; Ixodidae; Amblyomma.
REFERENCE	1 (bases 1 to 718)
AUTHORS	Nene, V., Lee, D., Quackenbush, J., Skilton, R., Mwaura, S., Gardner, M.J. and Bishop, R.
TITLE	AvGI, an index of Genes transcribed in the salivary glands of the ixodid tick Amblyomma variegatum
JOURNAL	Int. J. Parasitol. 32 (12), 1447-1456 (2002)
MEDLINE	2281296
PUBMED	12392910
COMMENT	Contact: Vish Nene Parasite Genomics Group The Institute for Genomic Research 9712 Medical Center Drive, Rockville, MD 20850, USA Tel: 301-610-5968 Fax: 301-838-0208 Email: nene@tigr.org Seq primer: M13 reverse.
LOCATION/QUALIFIERS	Location/Qualifiers 1. .594 /organism="Amblyomma variegatum" /mol_type="mRNA" /db_xref="taxon:34610" /clone="AVAA069" /tissue_type="Salivary glands" /dev_stage="Adult" /lab_host="E.coli strain DH10B-TonA" /clone_lib="AvSG" /note="Vector: pCMV-SPORT6.1; Salivary glands were dissected on day five after initiation of feeding. Total RNA was prepared using acid guanidium thiocyanate-phenol-chloroform extraction. The cDNA library was custom prepared by Invitrogen Corporation. Briefly, first strand cDNA was primed using oligo(dT) containing a NotI site. Size fractionated double stranded cDNA was ligated to EcoRV-NotI cleaved vector and electroporated into E.coli."
FEATURES	source
PRIMER	Seq Primer: M13 reverse.
LOCATION/QUALIFIERS	Location/Qualifiers 1. .718 /organism="Amblyomma variegatum" /mol_type="mRNA" /db_xref="taxon:34610" /clone="AVAA069" /tissue_type="Salivary glands" /dev_stage="Adult" /lab_host="E.coli strain DH10B-TonA" /clone_lib="AvSG" /note="Vector: pCMV-SPORT6.1; Salivary glands were dissected on day five after initiation of feeding. Total RNA was prepared using acid guanidium thiocyanate-phenol-chloroform extraction. The cDNA library was custom prepared by Invitrogen Corporation. Briefly, first strand cDNA was primed using oligo(dT) containing a NotI site. Size fractionated double stranded cDNA was ligated to EcoRV-NotI cleaved vector and electroporated into E.coli."
ORIGIN	
ALIGNMENT SCORES:	
Y Y	Score: 4.66e-47 Qy 1 LysLysThrThrArgArgAsnAlaTrpGlyAsnNetSerTyraGluLeuIleThr 20 Db 307 AAAAGAACCTGCTGCCATTTCAGGCAAGGGCACAGCAACAGCTCCGTGCAAGCCCTAGGCGAACATGCTCTAGGCCACAGTACAGTGATGACCTGATACG 366 b b Percent Similarity: 93.07% Best Local Similarity: 83.17% Query Match: 85.59% DB: 12 Length: 718 Matches: 84 Conservative: 10 Mismatches: 7 Index: 0 Gaps: 0
Y Y	21 ThrAlaIleMetAlaSerProGluIysArgLeuThrLeuIleGlnValTygGlutrpMet 40 Qy 21 ThrAlaIleMetAlaSerProGluIysArgLeuThrLeuIleGlnValTygGlutrpMet 40 Db 367 CAAGCUCATCAAAGCGCCGAGAACGCTCCAGAACATGCTCTAGGCCACAGTACAGTGATGACCTGATACG 426 b b Percent Similarity: 93.07% Best Local Similarity: 83.17% Query Match: 85.59% DB: 12 Length: 718 Matches: 84 Conservative: 10 Mismatches: 7 Index: 0 Gaps: 0
Y Y	41 ValGlnAsnValProTrpPheArgAspSerAsnSerAlaGlyTrpLys 60 Qy 41 ValGlnAsnValProTrpPheArgAspSerAsnSerAlaGlyTrpLys 60 Db 427 GTGCAGAACGCTGCCATTTCAGGCAAGGGCACAGCAACAGCTCCGTGCAAGCCCTAGGCGAACATGCTCTAGGCCACAGTACAGTGATGACCTGATACG 486 b b Percent Similarity: 93.07% Best Local Similarity: 83.17% Query Match: 85.59% DB: 12 Length: 718 Matches: 84 Conservative: 10 Mismatches: 7 Index: 0 Gaps: 0
Y Y	61 AsnSerIleArgHisAsnLeuSerLeuIleHisSerArgPheMetAsnProArg 100 Qy 61 AsnSerIleArgHisAsnLeuSerLeuIleHisSerArgPheMetAsnProArg 100 Db 487 AACTCTATCCGGCACACCTGCTGCAAGCATGCGATTCAGGCAAGCATGCTCTAGGCCACAGTACAGTGATGACCTGATACG 546 b b Percent Similarity: 93.07% Best Local Similarity: 83.17% Query Match: 85.59% DB: 12 Length: 718 Matches: 84 Conservative: 10 Mismatches: 7 Index: 0 Gaps: 0
Y Y	81 AlaGlyLysSerSerTrpTrpValIleAsnProAspAlaSerGlnGluGly 80 Qy 81 AlaGlyLysSerSerTrpTrpValIleAsnProAspAlaSerGlnGluGly 80 Db 81 AlaGlyLysSerSerTrpTrpValIleAsnProAspAlaSerGlnGluGly 80 b b Percent Similarity: 93.07% Best Local Similarity: 83.17% Query Match: 85.59% DB: 12 Length: 718 Matches: 84 Conservative: 10 Mismatches: 7 Index: 0 Gaps: 0
Y Y	101 Arg 101 Qy 101 Arg 101 Db 101 Arg 101 b b Percent Similarity: 93.07% Best Local Similarity: 83.17% Query Match: 85.59% DB: 12 Length: 718 Matches: 84 Conservative: 10 Mismatches: 7 Index: 0 Gaps: 0
DEFINITION	EST577137 AvSG Amblyomma variegatum cDNA clone AVAA069 5' end, mRNA sequence.
ACCESSION	BM290603
VERSION	1
KEYWORDS	EST, Amblyomma variegatum
SOURCE	Organism: Amblyomma variegatum Eukaryota: Metazoa; Arthropoda; Chelicera; Arachnida; Acari;
ORGANISM	Parasitiformes; Ixodida; Ixodidae; Amblyomma.
REFERENCE	1 (bases 1 to 718)
AUTHORS	Nene, V., Lee, D., Quackenbush, J., Skilton, R., Mwaura, S., Gardner, M.J. and Bishop, R.
TITLE	AvGI, an index of Genes transcribed in the salivary glands of the ixodid tick Amblyomma variegatum
JOURNAL	Int. J. Parasitol. 32 (12), 1447-1456 (2002)
MEDLINE	2281296
PUBMED	12392910
COMMENT	Contact: Vish Nene Parasite Genomics Group The Institute for Genomic Research 9712 Medical Center Drive, Rockville, MD 20850, USA Tel: 301-610-5968 Fax: 301-838-0208 Email: nene@tigr.org Seq primer: M13 reverse.
LOCATION/QUALIFIERS	Location/Qualifiers 1. .594 /organism="Amblyomma variegatum" /mol_type="mRNA" /db_xref="taxon:34610" /clone="AVAA069" /tissue_type="Salivary glands" /dev_stage="Adult" /lab_host="E.coli strain DH10B-TonA" /clone_lib="AvSG" /note="Vector: pCMV-SPORT6.1; Salivary glands were dissected on day five after initiation of feeding. Total RNA was prepared using acid guanidium thiocyanate-phenol-chloroform extraction. The cDNA library was custom prepared by Invitrogen Corporation. Briefly, first strand cDNA was primed using oligo(dT) containing a NotI site. Size fractionated double stranded cDNA was ligated to EcoRV-NotI cleaved vector and electroporated into E.coli."
FEATURES	source
PRIMER	Seq Primer: M13 reverse.
LOCATION/QUALIFIERS	Location/Qualifiers 1. .718 /organism="Amblyomma variegatum" /mol_type="mRNA" /db_xref="taxon:34610" /clone="AVAA069" /tissue_type="Salivary glands" /dev_stage="Adult" /lab_host="E.coli strain DH10B-TonA" /clone_lib="AvSG" /note="Vector: pCMV-SPORT6.1; Salivary glands were dissected on day five after initiation of feeding. Total RNA was prepared using acid guanidium thiocyanate-phenol-chloroform extraction. The cDNA library was custom prepared by Invitrogen Corporation. Briefly, first strand cDNA was primed using oligo(dT) containing a NotI site. Size fractionated double stranded cDNA was ligated to EcoRV-NotI cleaved vector and electroporated into E.coli."
ORIGIN	
ALIGNMENT SCORES:	
Y Y	Score: 4.66e-47 Qy 1 LysLysThrThrArgArgAsnAlaTrpGlyAsnNetSerTyraGluLeuIleThr 20 Db 307 AAAAGAACCTGCTGCCATTTCAGGCAAGGGCACAGCAACAGCTCCGTGCAAGCCCTAGGCGAACATGCTCTAGGCCACAGTACAGTGATGACCTGATACG 366 b b Percent Similarity: 93.07% Best Local Similarity: 83.17% Query Match: 85.59% DB: 12 Length: 718 Matches: 84 Conservative: 10 Mismatches: 7 Index: 0 Gaps: 0
Y Y	21 ThrAlaIleMetAlaSerProGluIysArgLeuThrLeuIleGlnValTygGlutrpMet 40 Qy 21 ThrAlaIleMetAlaSerProGluIysArgLeuThrLeuIleGlnValTygGlutrpMet 40 Db 367 CAAGCUCATCAAAGCGCCGAGAACGCTCCAGAACATGCTCTAGGCCACAGTACAGTGATGACCTGATACG 426 b b Percent Similarity: 93.07% Best Local Similarity: 83.17% Query Match: 85.59% DB: 12 Length: 718 Matches: 84 Conservative: 10 Mismatches: 7 Index: 0 Gaps: 0
Y Y	41 ValGlnAsnValProTrpPheArgAspSerAsnSerAlaGlyTrpLys 60 Qy 41 ValGlnAsnValProTrpPheArgAspSerAsnSerAlaGlyTrpLys 60 Db 427 GTGCAGAACGCTGCCATTTCAGGCAAGGGCACAGCAACAGCTCCGTGCAAGCCCTAGGCGAACATGCTCTAGGCCACAGTACAGTGATGACCTGATACG 486 b b Percent Similarity: 93.07% Best Local Similarity: 83.17% Query Match: 85.59% DB: 12 Length: 718 Matches: 84 Conservative: 10 Mismatches: 7 Index: 0 Gaps: 0
Y Y	61 AsnSerIleArgHisAsnLeuSerLeuIleHisSerArgPheMetAsnProArg 100 Qy 61 AsnSerIleArgHisAsnLeuSerLeuIleHisSerArgPheMetAsnProArg 100 Db 61 AACTCTATCCGGCACACCTGCTGCAAGCATGCGATTCAGGCAAGCATGCTCTAGGCCACAGTACAGTGATGACCTGATACG 546 b b Percent Similarity: 93.07% Best Local Similarity: 83.17% Query Match: 85.59% DB: 12 Length: 718 Matches: 84 Conservative: 10 Mismatches: 7 Index: 0 Gaps: 0
Y Y	81 AlaGlyLysSerSerTrpTrpValIleAsnProAspAlaSerGlnGluGly 80 Qy 81 AlaGlyLysSerSerTrpTrpValIleAsnProAspAlaSerGlnGluGly 80 Db 81 AlaGlyLysSerSerTrpTrpValIleAsnProAspAlaSerGlnGluGly 80 b b Percent Similarity: 93.07% Best Local Similarity: 83.17% Query Match: 85.59% DB: 12 Length: 718 Matches: 84 Conservative: 10 Mismatches: 7 Index: 0 Gaps: 0
Y Y	101 Arg 101 Qy 101 Arg 101 Db 101 Arg 101 b b Percent Similarity: 93.07% Best Local Similarity: 83.17% Query Match: 85.59% DB: 12 Length: 718 Matches: 84 Conservative: 10 Mismatches: 7 Index: 0 Gaps: 0
DEFINITION	EST577137 AvSG Amblyomma variegatum cDNA clone AVAA069 5' end, mRNA sequence.
ACCESSION	BM290603
VERSION	1
KEYWORDS	EST, Amblyomma variegatum
SOURCE	Organism: Amblyomma variegatum Eukaryota: Metazoa; Arthropoda; Chelicera; Arachnida; Acari;
ORGANISM	Parasitiformes; Ixodida; Ixodidae; Amblyomma.
REFERENCE	1 (bases 1 to 718)
AUTHORS	Nene, V., Lee, D., Quackenbush, J., Skilton, R., Mwaura, S., Gardner, M.J. and Bishop, R.
TITLE	AvGI, an index of Genes transcribed in the salivary glands of the ixodid tick Amblyomma variegatum
JOURNAL	Int. J. Parasitol. 32 (12), 1447-1456 (2002)
MEDLINE	2281296
PUBMED	12392910
COMMENT	Contact: Vish Nene Parasite Genomics Group The Institute for Genomic Research 9712 Medical Center Drive, Rockville, MD 20850, USA Tel: 301-610-5968 Fax: 301-838-0208 Email: nene@tigr.org Seq primer: M13 reverse.
LOCATION/QUALIFIERS	Location/Qualifiers 1. .594 /organism="Amblyomma variegatum" /mol_type="mRNA" /db_xref="taxon:34610" /clone="AVAA069" /tissue_type="Salivary glands" /dev_stage="Adult" /lab_host="E.coli strain DH10B-TonA" /clone_lib="AvSG" /note="Vector: pCMV-SPORT6.1; Salivary glands were dissected on day five after initiation of feeding. Total RNA was prepared using acid guanidium thiocyanate-phenol-chloroform extraction. The cDNA library was custom prepared by Invitrogen Corporation. Briefly, first strand cDNA was primed using oligo(dT) containing a NotI site. Size fractionated double stranded cDNA was ligated to EcoRV-NotI cleaved vector and electroporated into E.coli."
FEATURES	source
PRIMER	Seq Primer: M13 reverse.
LOCATION/QUALIFIERS	Location/Qualifiers 1. .718 /organism="Amblyomma variegatum" /mol_type="mRNA" /db_xref="taxon:34610" /clone="AVAA069" /tissue_type="Salivary glands" /dev_stage="Adult" /lab_host="E.coli strain DH10B-TonA" /clone_lib="AvSG" /note="Vector: pCMV-SPORT6.1; Salivary glands were dissected on day five after initiation of feeding. Total RNA was prepared using acid guanidium thiocyanate-phenol-chloroform extraction. The cDNA library was custom prepared by Invitrogen Corporation. Briefly, first strand cDNA was primed using oligo(dT) containing a NotI site. Size fractionated double stranded cDNA was ligated to EcoRV-NotI cleaved vector and electroporated into E.coli."
ORIGIN	

Pred. No.: 1.B4e-46 Length: 590
 Score: 469.00 Matches: 83
 Percent Similarity: 93.07% Conservative: 1.1
 Best Local Similarity: 82.18% Mismatches: 7
 Query Match: 84.50% Indels: 0
 DB: 12 Gaps: 0

US-09-844-353a-54 (1-103) x BM291076 (1-590)

QY 1 LysLysThrThrArgArgAsnAlaTyrGlyAsnMetSerTyrAlaGluLeuIleThr 20
 Db 255 AAAAGAACAGTCGCTGGCAACCTGGGCACATGTCTAGGGACCTGTTACG 314

QY 21 ThrAlaIleMetAlaSerProGluLysArgLeuIleAlaGlnValTyrGluTrpMet 40
 Db 315 CAAGCCATCCAAAGGCCGGAGAAAGCCTCACCGTGTACAGTAGTGATG 374

QY 41 ValGlnAsnValProTyrrPheArgAspLysGlyAspSerSerSerAlaGlyTrpLys 60
 Db 375 GTGCCAGAACCTGGCCCTATTCAAGGACAAGGGGACAGCAACGCTCCGGCTGGAG 434

QY 61 AsnSerIleArgHisAsnLeuSerLeuHisserArgPheMetArgTleGlnAsnGluGly 80
 Db 435 AACTCTATCCGGCACACCTGTCACAGGAGATTCATGGGAGTCCAGAAGGGC 494

QY 81 AlaGlyLysSerSerTrpTrpValLeAsnProAspAlaLysProGlyMetAsnProArg 100
 Db 495 GCGGGGAGAGCTGTTGATGCAACGGGACAAAGCCGGAAACCGGAGTCAACGG 554

QY 101 Arg 101
 Db 555 CGC 557

RESULT 5
 BM291800 715 bp mRNA linear EST 01-JUL-2002
 DEFINITION EST74342 AvSG Amblyomma variegatum cDNA clone AVABC60 5' end, mRNA sequence.
 ACCESSION BM291800
 VERSION 1
 KEYWORDS EST
 ORGANISM Amblyomma variegatum
 PARASITIFORMES: Ixodidae; Ixodidae; Amblyomma.
 Eukaryote; Metazoa; Arthropoda; Cheliceraata; Arachnida; Acari;
 1 (bases 1 to 715)
 Nene, V., Lee, D., Shackenbush, J., Skilton, R., Mwaura, S.,
 Gardner, M.J., and Bishop, R.
 TITLE AvGI, an index of genes transcribed in the salivary glands of the
 ixodid tick Amblyomma variegatum
 JOURNAL Int. J. Parasitol. 32 (12), 1447-1456 (2002)
 MEDLINE 22281296
 PUBMED 12392910
 COMMENT Contact: Vish Nene
 Parasite Genomics Group
 The Institute for Genomic Research
 9712 Medical Center Drive, Rockville, MD 20850, USA
 Tel: 301-611-5968
 Fax: 301-338-0208
 Seq Primer: M13 reverse.

FEATURES source
 Location/Qualifiers 1..715
 /organism="Amblyomma variegatum"
 /mol_type="mRNA"
 /db_xref="Taxon:34610"
 /clone="AVABC60"
 /tissue_type="Salivary glands"
 /dev_stage="Adult"
 /lab_host="E.coli strain DH10B-TonA"
 /clone_lib="AvSG"
 /note="Vector: pCMV-SPORT6.1; Salivary glands were dissected on day five after initiation of feeding. Total

FEATURES

RNA was prepared using acid guanidium thiocyanate-phenol-chloroform extraction. The cDNA library was custom prepared by Invitrogen Corporation. Briefly, first strand cDNA was primed using oligo(dT) containing a NotI site. Size fractionated double stranded cDNA was ligated to EcoRV-NotI cleaved vector and electroporated into E.coli."

ORIGIN

Alignment Scores:
 Pres. No.: 2.17e-44
 Score: 453.00
 Percent Similarity: 93.75%
 Best Local Similarity: 82.29%
 Query Match: 81.62%
 DB: 12 Gaps: 0
 US-09-844-353a-54 (1-103) x BM291076 (1-715)
 QY 1 LysLysThrThrArgArgAsnAlaTyrGlyAsnMetSerTyrAlaGluLeuIleThr 20
 Db 260 AAAAGAACACTGTCGCTGGCAACCTGGGCACATGTCCTAGGGACCTGTTACG 319
 QY 21 ThraIleIleMetAlaSerProGluLysArgLeuIleAlaGlnValTyrGluTrpMet 40
 Db 320 CAAGCCATCCAAAGGCCGGAGAAAGCCTCACCGTGTACAGTAGTGATG 379
 QY 41 ValGlnAsnValProTyrrPheArgAspLysGlyAspSerSerSerAlaGlyTrpLys 60
 Db 380 GTGCAGAACGTCGCCATTTCAGAACAGGAGAAGGAGAAG 439
 QY 61 AsnSerIleArgHisAsnLeuSerLeuHisSerArgPheMetAsnGluGly 80
 Db 440 AACCTCATCCAAAGGCCGGAGAAACGGGAGTCAACCTGCAACGTAACG 499
 QY 81 AlAglLysSerSerTrpTrpValLeAsnProAspAlaLysProGly 96
 Db 500 GCCGGGAGAGCTGTTGATGCAACGGGAGTCAACCTGCAACGTAACG 547
 RESULT 6
 BX673604
 LOCUS BX673604 Sus scrofa library 593 bp mRNA linear EST 28-OCT-2003
 DEFINITION BX673604 Sus scrofa cDNA clone scac0371.g.15 5prim, mRNA sequence.
 ACCESSION BX673604
 VERSION 1
 KEYWORDS EST
 SOURCE Sus scrofa (pig)
 ORGANISM Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Cetartiodactyla; Suina; Suidae; Sus.
 REFERENCE 1 (bases 1 to 593)
 AUTHORS Bonnet,A., Tosser-Klopp,G., Benne,F., Cabau,C., Villeger,S., Soares,M., Bonaldo,P. and Hately,F.
 TITLE A pig Normalized Multi-Tissue cDNA Library
 JOURNAL Unpublished (2003)
 COMMENT Contact: Tosser-Klopp G
 Institut National de la Recherche Agronomique
 Chemin de Borda-Rouge - Auzeville BP27, 31326 Castanet-Tolosan
 cedex, FRANCE
 Tel: 33 (0) 5.61.28.51.14
 Fax: 33 (0) 5.61.28.53.08
 Email: tosser@oulouse.inra.fr
 Clone distribution: AGENEA Resource centre, Francois Piumi, CEA Radiobiologie et Etude du
 Genome (IREG), Domaine de Vilvert, 78352, Jouy-en-Josas cedex,
 FRANCE, +33 (0) 1.34.65.28.02, +33 (0) 1.34.65.22.73
 Sequence cleaned of vector, adaptor and repetitions. Contact us at sgensupport@jouy.inra.fr to obtain the chromatogram of this sequence.
 Plate: 0037, row: 9 column: 15.
 Location/Qualifiers Location/Qualifiers

CDNA Library Preparation: Life Technologies, Inc.
 CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL)
 DNA Sequencing by: Incyte Genomics, Inc.
 Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LNL at: <http://image.llnl.gov>
 Plate: LLNAM679 row: b column: 05
 High quality sequence stop: 628.

FEATURES

source

- /organism="Homo sapiens"
- /mol_type="mRNA"
- /db_xref="taxon:9606"
- /clone="IMAGB:389251-6"
- /tissue_type="large cell carcinoma, undifferentiated"
- /lab_host="DH10B (phage-resistant)"
- /clone_lib="NIH MCC 69"
- /note="Organ: lung; Vector: pCMV-SPORT6; Site_1: NotI; Site_2: SalI; Cloned unidirectionally. Primer: Oligo dT. Average insert size 1.1 kb. Library constructed by Life Technologies."

ORIGIN

Alignment Scores:

Pred. No.:	Length:	Score:	Best Local Similarity:	Query Match:	DB:
5.84e-43	628	90.20%	90.45%	79.37%	10

US-09-844-353A-54 (1-103) × BB881329 (1-628)

Qy 1 LysLysThrThrThrArgGasnAlaTrpGlyAsnMetSerTyAlaGluLeuIleThr 20
 Db 206 AGGAATGTTCTGGTCTGGGGAAACGGCTGCTAACGGGACCTGTACCC 265

Qy 21 ThrAlaIleMetAlaSerProLeuLysArgLeuIleLeuAlaGlnValTyrglutrpMet 40
 Db 266 CGCGCCATCGAGCTGCCGACAAACGGCTACTGTCCTGATCTAGAGTGATG 325

Qy 41 ValGlnAsnValProTyrrpheArgAspLysGlyAspSerAsnSerAlaGlyTrpLys 60
 Db 326 GTGCGTGTGCGTCCCTAGCTTAAGGAAAGGGCAAGCAAAGCTGTGCCGGCTGAAG 385

Qy 61 AsnSerIleArchIleSerIleHisSerArgPheMetArgLeuGlnAsnGluGly 80
 Db 386 AACTCCATCCGGCACACCTGTACATCTGATCATCTGATTCAGGGCAATGAGGGA 445

Qy 81 AlaGlyLysSerSerTrpValIleAsnProAsp --AlaLysProGlyMetAsnPro 99
 Db 446 ACTGGCAAGAGCTCTGGATCATCAACCTGTATGGATCAGGGAAAAGGCC 505

Qy 100 ArgArg 101
 Db 506 CGCGGG 511

RESULT 9

BC019532 BC019532 998 bp mRNA linear HTC 20-SEP-2002

LOCUS Mus musculus, Similar to forehead box 03, clone IMAGE:4162687, mRNA.

DEFINITION BC019532.1 GI:18044075

VERSION 1

KEYWORDS HTC.

SOURCE Mus musculus (house mouse)

ORGANISM Eukaryota; Metazoa; Chordata; Craniata: Vertebrata; Buteostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus. (bases 1 to 998)

REFERENCE Strauberg, R.

TITLE Direct Submission

JOURNAL Submitted (19-DEC-2001) National Institutes of Health, Mammalian

Gene Collection (MGC), Cancer Genomics Office, National Cancer Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590, USA

NIH-MGC Project URL: <http://mgc.nci.nih.gov>

COMMENT Contact: MGC help desk
 Email: cgbps@mail.nih.gov
 Tissue Procurement: Jeffrey E. Green, M.D.
 CDNA Library Preparation: Life Technologies, Inc.
 CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL)
 DNA Sequencing by: Baylor College of Medicine Human Genome Sequencing Center
 Center code: BCM-HGSC
 Web site: <http://www.hgsc.bcm.tmc.edu/cDNA/>
 Contact: amg@bcm.tmc.edu
 Gunarane, P.H.; Garcia, A.M.; Lu, X.; Hulyk, S.W.; Hale, S.M.; Yoon, V.S.; Kowis, C.R.; Lawrence, S.; Martin, R.G.; Muzny, D.M.; Richards, S.; Gibbs, R.A.

Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LNL at: <http://image.llnl.gov>

Series: IRAK Plate: 37 Row: d Column: 23

This clone was selected for full length sequencing because it passed the following selection criteria: matched mRNA gi: 9789950 This clone has the following problem: frame shifted.

FEATURES

Source

- 1. /organism="Mus musculus"
- /mol_type="mRNA"
- /strain="FVB/N"
- /db_xref="taxon:10090"
- /clone="IMAGE:4162687"
- /tissue_type="Liver, normal, 5 month old male mouse."
- /clone_id="NCI CGAP Li9"
- /lab_host="DH10B"
- /note="Vector: pCMV-SPORT6"

ORIGIN

Alignment Scores:

Pred. No.:	Length:	Score:	Best Local Similarity:	Query Match:	DB:
1.2e-42	998	44.50%	90.20%	77.45%	DB:

Qy 1 LysLysThrThrThrArgGasnAlaTrpGlyAsnMetSerTyAlaGluLeuIleThr 20
 Db 604 AGGAATGCTCTGGGGAAATGCTGGGGAACTGTCCTATGCCGACCTGATCACC 663

Qy 21 ThrAlaIleMetAlaSerProLeuLysArgLeuIleLeuAlaGlnValTyrglutrpMet 40
 Db 664 CGGCCATCGAGCTGCCCTGCGGGAACTGTCCTATGCCGACCTGATCAGGATGATG 723

Qy 41 ValGlnAsnValProTyrrpheArgAspLysGlyAspSerAsnSerAlaGlyTrpLys 60
 Db 724 GTCGCTGTGCGCCFACTTAAGGATAAGGGACAGCAAGCTGCTGCGCTGAG 783

Qy 61 AsnSerIleArchIleSerIleHisSerArgPheMetArgLeuGlnAsnGluGly 80
 Db 784 AACATCCATCCGGCACACCTGTCCTGCCCTCATGCCGCTTACAGGATGAGGC 843

Qy 81 AlaGlyLysSerSerTrpValIleAsnProAsp --AlaLysProGlyMetAsnPro 99
 Db 844 ACGGCAAGAGCTCTGGATCATGACCCGATGGATGAGG 903

Qy 100 ArgArg 101
 Db 904 CGCGGG 909

RESULT 10

AK047413

1.101192-11.01959, 1.1217928-1.1220615); NCI CGAP Co10 pool 1
 LIAM 2644-2653, 2871-2872 (IMAGE ClonetIDs 1057416-1061255,
 1144584-1145351). Subtraction was performed as previously
 described [Bonaldo, Lennon & Soares (1996): Normalization
 and Subtraction: Two Approaches To Facilitate Gene
 Discovery. Genome Research 6, 791-806.
 TAG_TISSUE-colon
 TAG_LIB=NCI CGAP Co10
 TAG_SSQ=AAAAGC"

ORIGIN

Alignment Scores:	
Pred. No.:	Length:
Score:	427.50
Percent Similarity:	91.18%
Best Local Similarity:	73.53%
Query Match:	77.03%
DB:	10
Gaps:	1

US-09-844-353A-54 (1-103) × AW204560 (1-459)

Qy	Db	Score:	Length:
1 LysLysThrThrThrArgArgAsnAlaTrpGlyAsnMetSerTyrAlaGluLeuIleThr 20	153 AAAGAGAGCTGCTCCGCCCGAACGGCTGGSOCACCTGCTCACCGACTCATCACC 212	427.50	459
21 ThralalineMetAlaSerProGluLysArgLeuThrIleuAlaGlnValtyrGluLtrPMet 40	213 AAGGCCATCGAGCTGGAGCTGGCGAGAACGGGTCAACCTGTGCAAGTCATACGAATGTGATG 272	91.18%	75
4 ValGlnAsnValProLysPheArgArgLysGlyAsnSerIleSerAlaGlyTrPlys 60	273 GTCAAGCTGGCCCTACTCTAACGGTAAGCTGACAGAACAGCTGGGGCTGGAG 332	73.53%	18
61 AsnSerIleArgHisAsnLeuSerLeuHisserArgPheMetArgIleGinAsnGluGly 80	333 AATTCAATTCTCTCATATCTCTTCAAGCACTTCATCGCAACTTCATCGTGCGAAATGAGA 392	77.03%	8
81 AlaGlyLysSerSerPtpTrpValleAsnProASP---AlaLysProGlyMetAsnPro 99	393 ACTGGAAAAAGTTCTGGATGCTCAATCCAGGGTCCAAGAGGGAAATCCCT 452	77.03%	8
100 ArgArg 101			
Db	453 AGGAGA 458		

RESULT 13
 BI181558 LOCUS BI181558 492 bp mRNA linear EST 10-JUL-2001
 DEFINITION UNI-P-FN-ao-e-07-0-UNI-P-FN-S1 UNI-P-FN S1 mRNA clone
 ACCESSION BI181558 1 mRNA sequence.
 VERSION BI181558.1 GI:14655967
 EST.
 KEYWORDS SOURCE Sus scrofa (pig)
 ORGANISM Sus scrofa
 Mammalia; Eutheria; Chordata; Craniata; Vertebrata; Euteleostomi;
 Caetano,A.R., Johnson,R.K. and Pomp,D.
 AUTHORS Generation and sequence characterization of a normalized cDNA
 TITLE Library from swine ovarian follicles
 JOURNAL Mamm. Genome 14 (1), 65-70 (2003)
 MEDLINE 22419904
 PUBMED 12532269
 COMMENT Contact: Pomp, D
 Department of Animal Science
 University of Nebraska, Lincoln
 Lincoln, NE 68583-0908, USA
 Tel: 402 472 6416
 Fax: 402 472 3362
 Email: dpomp@unl.edu
 Oligo-dB track not found. Not 1 site shown in beginning of sequence
 is likely internal to the message. The following repetitive

ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Butheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 505)
REFERENCE NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.

AUTHORS National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
Tumor Gene Index

JOURNAL Unpublished (1997)

COMMENT Contact: Robert Straubberg, Ph.D.
Email: cga@rmail.nih.gov
This clone is available royalty-free through LInL contact the
IMAGE Consortium (infoimage.lnl.gov) for further information.

Insert Length: 875 Std Error: 0.00
Seq primer: -40UP from Gibco
High quality sequence stop: 466.

FEATURES Location/Qualifiers

source 1..505
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:2062248"
/tissue_type="pooled human melanocyte, fetal heart, and
pregnant uterus"
/lab_host="Soares NHMPU S1"
/clone_lib="Soares NHMPU S1"
/note="Organ: mixed (see below); Vector: pMT3D-Pac
(Pharmacia) with a modified polylinker. Site 1: Not I;
Site 2: Eco RI; Equal amounts of plasmid DNA from three
normalized libraries (melanocyte 2NBHM, pregnant uterus
NbHPV, and fetal heart NbHH19W) were mixed, and ss circles
were made in vitro. Following HAP purification, this DNA
was used as tracer in a subtractive hybridization
reaction. The driver was PCR-amplified cDNAs from pools of
5,000 clones made from the same 3 libraries. The pools
consisted of I.M.A.G.E. clones 260232-265223,
340488-345479, and 484488-485479."

ORIGIN

Alignment Scores:

Prod. No.:	Length:	Matches:	Score:	Percent Similarity:	Best Local Similarity:	Query Match:	DB:
1.58E-41	505	75	427.50	91.18%	73.53%	77.03%	

QY 1 LysLysThrThrArgArgAsnAlaTrpGlyAsnMetSerTyrAlaGluLeuIleThr 20
Db 1.54 ARGAGCCTCCGCCGGCAACCTGTCTACGGCACTCTCATACC 213

QY 21 ThrAlaIleMetAlaSerProGluLysArgLeuThrLeuAlaGlnValTyrGluTrpMet 40
Db 214 AAGGCCATCGAGAGCTGGCGAGAAGCGGTACGGTGTGCCAGATCTAGAGTCATG 273

QY 41 ValGlnAsnValProTyrrPheArgAspLysGlyAspSerAsnSerSerAlaGlyTrpLys 60
Db 274 GTCAAAGCGTGCGCTACTTCAGATAAGGTGACGCCAACAGCTGGCAAG 333

QY 61 AsnSerIleArgHisAsnLeuSerLeuSerLeuSerArgPheMetArgIleGlnAsnGluGly 80
Db 334 AATTCAATTGCGCATATCTGCCCCGAACTGAGATGAGGA 393

QY 81 AlaGlyLysSerSerTrpTrpValIleAsnProAsp---AlaLysProGlyMetAsnPro 99
Db 394 ACTGGAAGAAAGTCTTCAGGATGCTCAATCCAGGGCAAGACGGGAATCTCC 453

QY 100 ArgArg 101
Db 454 AGGAGA 459

LOCUS AI382221 NCI CGAP CBL1 Homo sapiens cDNA clone IMAGE:2074664 3'
DEFINITION similar to gb:U02368 PAIR ED BOX PROTEIN PAX-3 (HUMAN) ; mrNA
sequence.

ACCESSION AI382221
VERSION AI382221.1 GI:4195002
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Mammalia: Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE NCBI-NCAP http://www.ncbi.nlm.nih.gov/ncicgap.
AUTHORS National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
TITLE Tumor Gene Index
JOURNAL Unpublished (1997)
COMMENT Contact: Robert Straubberg, Ph.D.
Email: cga@rmail.nih.gov
Tissue Procurement: Ash Alizadeh, John Byrd, M.D., Mike Grever,
M.D., Louis M. Straud, M.D., Ph.D.
cDNA Library Preparation: M. Bento Soares, Ph.D.
DNA Sequencing by: Washington University Genome Sequencing Center
Clone distribution: NCI-NCAP clone distribution information can be
found through the I.M.A.G.E. Consortium/LInL at:
www-bio.lnl.gov/bbrp/image.html
Insert Length: 513 Std Error: 0.00
Seq primer: -40UP from Gibco
High quality sequence stop: 445.
FEATURES source
1..545
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:2074664"
/tissue_type="IB-cell", chronic lymphocytic leukemia"
/lab_host="DH10B"
/clone.lib="NCI CGAP CBL1"
/note="Vector: pMT3D-Pac (Pharmacia) with a modified
polylinker; Site 1: Not I; Site 2: Eco RI; 1st strand cDNA
was primed with a Not I - oligo(dT) primer [5',
TGTATCACATCTGAGTCGGAGCGCGATGCTTTTTTTTTTTTTTT
T 3'] ; double-stranded cDNA was ligated to Eco RI
adapters (Pharmacia), digested with Not I and cloned into
the Not I and Eco RI sites of the modified pMT3 vector.
Library is normalized, and was constructed by Bento
Soares and M.Fatima Bonaldo."

ORIGIN

Alignment Scores:

Prod. No.:	Length:	Matches:	Score:	Percent Similarity:	Best Local Similarity:	Query Match:	DB:
1.78E-41	545	75	427.50	91.18%	73.53%	77.03%	

QY 1 LysLysThrThrArgArgAsnAlaTrpGlyAsnMetSerTyrAlaGluLeuIleThr 20
Db 1.54 ARGAGCCTCCGCCGGCAACCTGTCTACGGCACTCTCATACC 213

QY 21 ThrAlaIleMetAlaSerProGluLysArgLeuThrLeuAlaGlnValTyrGluTrpMet 40
Db 214 AAGGCCATCGAGAGCTGGCGAGAAGCGGTACGGTGTGCCAGATCTAGAGTCATG 273

QY 41 ValGlnAsnValProTyrrPheArgAspLysGlyAspSerAsnSerSerAlaGlyTrpLys 60
Db 274 GTCAAAGCGTGCGCTACTTCAGATAAGGTGACGCCAACAGCTGGCAAG 333

QY 61 AsnSerIleArgHisAsnLeuSerLeuSerLeuSerArgPheMetArgIleGlnAsnGluGly 80
Db 334 AATTCAATTGCGCATATCTGCCCCGAACTGAGATGAGGA 393

QY 81 AlaGlyLysSerSerTrpTrpValIleAsnProAsp---AlaLysProGlyMetAsnPro 99
Db 394 ACTGGAAGAAAGTCTTCAGGATGCTCAATCCAGGGCAAGACGGGAATCTCC 453

QY 100 ArgArg 101
Db 454 AGGAGA 459

RESULT 15

Db 334 AATTCAATTCCTCATTAATCTGTCCCTACACGCAAGTTATTCTGCTGAGAATCAGGA 393
Qy 81 AlaGlyLysSerSerTrpTrpValleAsnProAsp--AlaLysProGlyMetAsnPro 99
Db 394 ACTGGAAAAGTTCTTGGTGGATGCTCAATCCAGGGTGGCAAGGCCAAAGGGAAATCTCT 453
Qy 100 ArgArg 101
Db 454 AGGAGA 459

Search completed: July 29, 2004, 12:13:31
Job time : 2496 secs

This Page Blank (uspto)

Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on: July 29, 2004, 09:52:26 ; Search time 3081 Seconds
 (without alignments)
 1448.988 Million cell updates/sec

Title: US-09-844-353A-54

Perfect score: 555

Sequence: 1 KKTTRRANQNMSTAILIT.....SSWWWINPDAKPGMMMPRRTR 103

Scoring table: BLOSUM62

Xgapop	10.0	Xgapext	0.5
Ygapop	10.0	Ygapext	0.5
Fgapop	6.0	Fgapext	7.0
Delop	6.0	Delext	7.0

Searched: 3470272 seqs, 21671516995 residues

Total number of hits satisfying chosen parameters: 6940544

Minimum DB seq length: 0

Maximum DB seq length: 20000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Command line parameters:

```
-MODEL=frame-p2n.model -DEV=xLP
-Q-/cgm2.1/OSPII spool.D/USI9844353/runat 23072004_084913_28237/app_query.fasta_1.263
-DE:GenImbl -QEMT:fastool -SUFFIXlarge -MINBATCH=0_1 -LOOPEXT=0
-UNITS:bits -START=1 -BND=-1 -MATRIX=blosun62 -TRANS:humana4.o.cgi -LIST=15
-DOCALIGN=200 -THR SCORE=pct -THR MAX=100 -THR MIN=0 -ALIGNN=15 -MODE=LOCAL
-OUTFMT=pto -NORMExt -HEAPSIZE=500 -MINLN=0 -MAXLN=20000000000
-USER=USI9844353 @CGN_1.1_7406 @runat 23072004_084913_28237 -NCPU=6 -ICPU=3
-NO_MMAPP -LARGEQUERY -NEG SCORE=0 -WAIT -DSBLOCK=100 -LONGLOG
-DEV TIMEOUT=120 -WARN TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7
```

Database : GenEmbl:*

1: gb_ba:*	7: gb_ph:*	13: gb_uni:*	19: em_mui:*	25: em_pl:*
2: gb_htg:*	8: gb_lv:*	14: gb_vl:*	20: em_om:*	26: em_ro:*
3: gb_in:*	9: gb_om:*	15: em_ba:*	21: em_or:*	27: em_sts:*
4: gb_mv:*	10: gb_ro:*	16: em_fun:*	22: em_ov:*	28: em_un:*
5: gb_sts:*	11: gb_pat:*	17: em_hum:*	23: em_pat:*	
6: gb_pat:*	12: gb_sy:*	18: em_in:*	24: em_ph:*	
7: gb_ph:*	13: gb_uni:*	19: em_mui:*	25: em_pl:*	
8: gb_lv:*	14: gb_vl:*	20: em_om:*	26: em_ro:*	
9: gb_in:*	15: em_ba:*	21: em_or:*	27: em_sts:*	
10: gb_ro:*	16: em_fun:*	22: em_ov:*	28: em_un:*	
11: gb_sts:*	17: em_hum:*	23: em_pat:*	24: em_ph:*	
12: gb_sy:*	18: em_in:*	19: em_mui:*	25: em_pl:*	
13: gb_uni:*	19: em_mui:*	20: em_om:*	26: em_ro:*	
14: gb_lv:*	20: em_om:*	21: em_or:*	27: em_sts:*	
15: em_ba:*	22: em_ov:*	23: em_pat:*	28: em_un:*	
16: em_fun:*	24: em_ph:*	25: em_pl:*		
17: em_hum:*	26: em_ro:*	27: em_sts:*		
18: em_in:*	28: em_un:*	28: em_un:*		

29: em_vi:*

30: em_htg_hum:*

31: em_htg_inv:*

32: em_htg_other:*

33: em_htg_mus:*

34: em_htg_fn:*

35: em_htg_rdn:*

36: em_htg_mam:*

37: em_htg_vrt:*

38: em_sy:*

39: em_htgo_hum:*

40: em_htgo_mus:*

41: em_htgo_other:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query	Match	Length	DB	ID	Description
1	555	100.0	3499	6	AR148563		AR148563 Sequence
2	555	100.0	3499	6	BD193386		BD193386 The rapate
3	549	98.9	2565	3	AF020343		AF020343 Caenorhab
4	549	98.9	2579	3	AF020342		AF020342 Caenorhab
5	549	98.9	3035	3	AF02112		AF02112
6	472	85.0	2589	3	AY281749		AY281749 Strongylo
7	450	81.1	450	3	AF46831		AF46831 Drosophil
8	450	81.1	3934	3	AY089542		AY089542 Drosophil
9	440.5	79.4	2685	9	AK092357		AK092357 Homo sapi
10	440.5	79.4	2870	9	AK122861		AK122861 Homo sapi
11	440.5	79.4	2889	10	AF114259		AF114259 Mus muscu
12	440.5	79.4	3183	9	AF032886		AF032886 Homo sapi
13	440.5	79.4	3300	9	BC020227		BC020227 Homo sapi
14	440.5	79.4	3300	9	BC021224		BC021224 Homo sapi
15	435.5	78.5	5742	5	AF114261		AF114261 Gallus ga
16	433.5	78.1	2740	5	AY040320		AY040320 Xiphophor
17	427.5	77.0	1968	9	BT007455		BT007455 Homo sapi
18	427.5	77.0	2162	4	AY094061		AY094061 Sus scrof
19	427.5	77.0	2413	9	BC021981		BC021981 Homo sapi
20	427.5	77.0	2503	10	AY255525		AY255525 Spermophil
21	427.5	77.0	3421	9	HSU02310		HSU02310 Human fork
22	427.5	77.0	3421	9	AX87664		AX87664 Sequence
23	427.5	77.0	5723	9	AF032885		AF032885 Homo sapi
24	427.5	77.0	5723	9	AX779990		AX779990 Sequence
25	426.5	76.8	5759	6	AF126056		AF126056 Mus muscu
26	426.5	76.8	2460	10	AF114258		AF114258 Mus muscu
27	426.5	76.8	4945	10	MMU252157		MMU252157 Mus muscu
28	425.5	76.7	4547	5	AF114262		AF114262 Danio rer
29	419.5	75.6	2528	10	MNU571630		MNU571630 Mus muscu
30	416.5	75.0	2103	10	AF126057		AF126057 Mus muscu
31	416.5	75.0	1719	9	AC016671		AC016671 Homo sapi
32	413.5	74.5	1733	10	AB032770		AB032770 Mus muscu
33	413.5	74.5	1993	10	AF114260		AF114260 Mus muscu
34	410.5	74.0	1073	6	AR242549		AR242549 Sequence
35	410.5	74.0	3394	6	AR338823		AR338823 Sequence
36	406.5	73.2	3171	6	AX821959		AX821959 Sequence
37	406.5	73.2	3171	9	HSAFX		HSAFX m
38	388.5	70.0	1722	9	HSAAJ1589		HSAAJ1589 Homo sapi
39	378.5	68.2	2704	6	AR148564		AR148564 Sequence
40	378.5	68.2	2704	6	BD193387		BD193387 Therapeut
41	378	68.1	900	5	AF064828		AF064828 Danio rer
42	372.5	67.1	2345	3	AF020344		AF020344 Caenorhab
43	356	64.1	1873	3	AY281750		AY281750 Strongylo
44	339.5	61.2	1000	10	AF247812		AF247812 Rattus no
45	335.5	60.5	1353	6	AX665875		AX665875 Sequence

ALIGNMENTS

AR148563	AR148563	3499 bp	DNA	linear	PAT 08-AUG-2001	COMMENT	THE GENERAL HOSPITAL CORP OS Cenorhabditis elegans
LOCUS	Sequence 43 from patent US 6225120.					PN	JP 2002511747-A/15
DEFINITION						PD	16-APR-2002
ACCESSION	AS148563					PF	15-MAY-1998 JP 1998549639
VERSION	AR148563 .1	GI:15112653				PR	15-MAY-1997 US 08/857076, 07-JUL-1997 US 08/88534 PI
KEYWORDS	Unknown.					GARY RUVKUN, KOTARO KIMURA, GARTH PATERSON, SCOTT OGG, SUZANNE PI	
SOURCE	Unknown.					PI PARADISE,	
ORGANISM	Unclassified.					PI HEIDI TISSENBAUM, JASON MORRIS, ALLISON KOEEK, SARAH PIERCE PC	
REFERENCE	1 (bases 1 to 3499) Ruvkun, G., Kimura, K., Patterson, G., Ogg, S., Paradis, S., Tissenbaum, H., Morris, J. and Kowek, A.					PI L1K49/00, C12N5/06, C07H21/04	
AUTHORS						CC Therapeutic and diagnostic tools for impaired glucose tolerance conditions	
TITLE	Therapeutic and diagnostic tools for impaired glucose tolerance conditions					FH Key	
JOURNAL	Patent: US 6225120-A 43 01-MAY-2001;					FT source	
FEATURES	Location/Qualifiers					Location/Qualifiers	
Source	1. 3499 /organism="Caenorhabditis elegans" /mol_type="genomic DNA" /db_xref="taxon:6239"					1. .3499 /organism="Caenorhabditis elegans" /mol_type="genomic DNA" /db_xref="taxon:6239"	
ORIGIN							
Alignment Scores:							
Pred. No. :	4.1e-52	Length:	3499			Alignment Scores:	
Score:	555.00	Matches:	103			Pred. No. :	4.1e-52
Percent Similarity:	100.00%	Conservative:	0			Score:	555.00
Best Local Similarity:	100.00%	Mismatches:	0			Percent Similarity:	100.00%
Query Match:	100.00%	Indels:	0			Best Local Similarity:	100.00%
DB:	6	Gaps:	0			Query Match:	100.00%
DB:						DB:	
US-09-844-353A-54 (1-103) × BD193386 (1-3499)							
Qy	1 LysLysThrThrThrArgArgAsnAlaTrpGlyAsnMetSerTyrAlaGluLeuIleThr 20	Qy	1 LysLythrThrThrArgArgAsnAlaTrpGlyAsnMetSerTyrAlaGluLeuIleThr 20				
Db	727 AAAAGCAACCAAGCGAACGGAACTGGAAATTGTCTATGACTTACT 786	Db	727 AAAAGCAACCAAGCGAACGGAACTGGAAATTGTCTATGACTTACT 786				
Qy	21 ThrAlaIleMetAlaSerProGlyLysArgLeuIleAlaGlnValtyGluTrpMet 40	Qy	21 ThrAlaIleMetAlaSerProGlyLysArgLeuIleAlaGlnValtyGluTrpMet 40				
Db	787 ACAGCCATTATGGCTATGGCTATGCTCAGAAACGGTAACCTTGCAACAAGTTACGAATGATG 846	Db	787 ACAGCCATTATGGCTATGGCTATGCTCAGAAACGGTAACCTTGCAACAAGTTACGAATGATG 846				
Qy	41 ValGlnAsnValProTyrrPheArgAspSerAsnSerSerAlaGlyTrpLys 60	Qy	41 ValGlnAsnValProTyrrPheArgAspSerAsnSerSerAlaGlyTrpLys 60				
Db	847 GTCCAGATGTTCCATACTTCAGGATAAGGAAACAGTTAGCTGGATGGAG 906	Db	847 GTCCAGATGTTCCATACTTCAGGATAAGGAAACAGTTAGCTGGATGGAG 906				
Qy	61 AsnSerIleArgHisAsnLeuSerLeuIleHisSerArgPheMetArgIleGlnAsnGluGly 80	Qy	61 AsnSerIleArgHisAsnLeuSerLeuIleHisSerArgPheMetArgIleGlnAsnGluGly 80				
Db	907 AACTGATCCGTACATCTGCTCTCATCGAANTCAGATGAGGA 966	Db	907 AACTGATCCGTACATCTGCTCTCATCGAANTCAGATGAGGA 966				
Qy	81 AlaGlyLysSerIleTrpTrpValileAsnProAspAlaLysProGlyMetAsnProArg 100	Qy	81 AlaGlyLysSerIleTrpTrpValileAsnProAspAlaLysProGlyMetAsnProArg 100				
Db	967 GCGGAAAGAGCTGTGGTTATATCCAGATGAAAGCCAGGATGATCCACGG 1026	Db	967 GCGGAAAGAGCTGTGGTTATATCCAGATGAAAGCCAGGATGATCCACGG 1026				
Qy	101 ArgThrArg 103	Qy	101 ArgThrArg 103				
Db	1027 CGTACACGTT 1035	Db	1027 CGTACACGTT 1035				
RESULT 2							
BD193386							
LOCUS	BD193386 3499 bp	DNA	linear	PAT 17-JUL-2003	RESULT 3	AF020343	
DEFINITION	Therapeutic and diagnostic tools for impaired glucose tolerance conditions.				LOCUS	AF020343	
ACCESSION	BD193386				DEFINITION	Caenorhabditis elegans	
VERSION	BD193386 .1	GI:33003125			ACCESSION	DAF-16a2 (daf-16) mRNA, complete cds.	
KEYWORDS	JP 2002511747-A/15.				VERSION	AF020343	
SOURCE	Caenorhabditis elegans				KEYWORDS	AF020343.1 GI:2618978	
ORGANISM	Bukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditoidea; Peledrinae; Caenorhabditis.				ORGANISM	Caenorhabditis elegans	
REFERENCE	1 (bases 1 to 3499) Ruvkun, G., Kimura, K., Patterson, G., Ogg, S., Paradis, S., Tissenbaum, H., Morris, J., Kowek, A. and Pierce, S.				REFERENCE	Rhabditoidea; Rhabditida; Peledrinae; Caenorhabditis.	
AUTHORS					AUTHORS	Ogg, S., Paradis, S., Gottlieb, S., Patterson, G. I., Lee, I., Tissenbaum, H. A. and Ruvkun, G.	
TITLE	Therapeutic and diagnostic tools for impaired glucose tolerance conditions				TITLE	The Fork head transcription factor DAF-16 transduces insulin-like metabolic and longevity signals in C. elegans	
JOURNAL	Patent: JP 2002511747-A 15 16-APR-2002;						

JOURNAL	Nature 389 (6654), 994-999 (1997)	LOCUS DEFINITION	AF020342 Caenorhabditis elegans fork head-related transcription factor DAF-16a1 (daf-16) mRNA, complete cds.
PUBMED	9303126	ACCESSION	AF020342
REFERENCE	2 (bases 1 to 2565)	VERSION	AF020342.1 GI:2618976
AUTHORS	Ogg,S., Paradis,S. and Ruvkun,G.	SOURCE	Caenorhabditis elegans
TITLE	Direct Submission	ORGANISM	Caenorhabditis elegans
JOURNAL	Submitted (21-AUG-1997) Molecular Biology, MGH, 50 Blossom St., Boston, MA 02114, USA	REFERENCE	Ogg,S., Paradis,S., Patterson,G.I., Lee,L., Tissenbaum,H.A., and Ruvkun,G.
FEATURES	Location/Qualifiers	AUTHORS	The Fork head transcription factor DAF-16 transduces insulin-like metabolic and longevity signals in C. elegans
source	1..565 /organism="Caenorhabditis elegans" /mol_type="mRNA" /strain="N" /db_xref="taxon:6239" /chromosome="1" /map="1; between mgp45 and mgp49" 1..2565 /gene="daf-16" /note="null phenotype is suppression of daf-2" 297..1823 /gene="daf-16" /codon_start=1 /product="fork head-related transcription factor DAF-16a2" /protein_id="AAAB8391_1" /db_xref="GI:2618979" /_translation="MMENLVQDGTDASSAATSTSSVSRFGADTFMNTPDDVMNDM EPIPLDRDNTWMPMRPQLEPPNNSPPIHOPERDLYGSENEQQCOLGARSSNST AMLHTPGDSNSHOTSESEMSSESDDVSGKRTTTRNAWMNSYALITTAIMSP KEFLTAQYEWNVQNVFPRKGDSNSAGWNSIREHNSRFRNIRIONGAQRSSW WTINPDAPKRPRTIERTNTETTKAQEKSRRGAKKRIKERAIMGSLSHISTLNG SIAGSICOTISHLYDDDSMGGFDNYDSSEFRPTQSNLISIPSSRSVSPAGSDTDD LEFPFWTGESVAIAPSIDVDRDQMIDATTGIVKQEVKQEVKPIKEPIAPPSTHE LNSVRGSAQNDLRLNPVPSNFNEKPKMLPQGAGYQNGTIPINNLSTSNSSPRLGI QSCTGIVAAQHTVASSALPIDENLTLPDQPLMDTVDALLRHELSQAGQHIFHDL	TITLE	The Fork head transcription factor DAF-16 transduces insulin-like metabolic and longevity signals in C. elegans
CDS	source	JOURNAL	Nature 389 (6654), 994-999 (1997) Molecular Biology, MGH, 50 Blossom St., Boston, MA 02114, USA
		FEATURES	Location/Qualifiers
		source	1..2579 /organism="Caenorhabditis elegans" /product="fork head-related transcription factor DAF-16a1" /strain="N" /db_xref="taxon:6239" /_chromosome="1" /map="1; between mgp45 and mgp49" 1..2579 /gene="daf-16" /note="null phenotype is suppression of daf-2" 299..1831
		CDS	
ORIGIN			
Alignment Scores:			
Prod. No.:	1.36e-51	Length:	2565
Score:	549.00	Matches:	102
Percent Similarity:	99.03%	Conservative:	0
Best Local Similarity:	99.03%	Mismatches:	1
Query Match:	98.92%	Indels:	0
DB:	3	Gaps:	0
US-09-844-353A-54 (1-103) x AF020343 (1-2565)			
Qy	1 LysLysThrThrArgArgAsnAlaTrpGlyAsnMetSerTyrAlaGluLeuIleThr	Alignment Scores:	1.37e-51
Db	693 AAAAGAACGACGACAGAACGAGAACGAGAACGAGAACCTATCACT	Pred. No.:	549.00
Qy	21 ThralAlleMetAlaSerIroGluLysArgLeuThrLeuAlaGlnValtryGluTrpMet	Score:	549.00
Db	753 ACAGCCATATGGCTAGTCAGAGAAA CGTTAACTCTTGACAAAGTTACGATGGATG	Percent Similarity:	99.03%
Qy	41 ValGlnAsnValProTyroPheArgAspLysGlyAspSerAsnSerAlaGlyTrpMet	Best Local Similarity:	99.03%
Db	813 GTCAGATGTCCATACTTCAGGATAGGGAGATGAACTGTTAGCTGGATGAAAG	Query Match:	98.92%
Qy	61 AsnSerIleArgHisAsnLeuSerLeuHisSerArgPheMetArgIleGlnAsnGluGly	DB:	3
Db	873 AACTCGATCGTCGTCACAATGCGATTCAGATGAGGA	ORIGIN	21 ThralAlleMetAlaSerProGlyLysArgLeuThrLeuAlaGlnValtryGluTrpMet
Qy	81 AlaGlyLysSerSerTrpValIleAsnProAspAlaLysProGlyMetAsnProArg	Qy	1 LysLysThrThrArgAsnAlaTrpGlyAsnMetSerTyrAlaGlnValtryGluTrpMet
Db	923 GCGGAAGAGCTGTCGGTTATTATCCAGATGCCAGAAAGGGATCCAGGG	Db	701 AAAAGAACGACGACAGAACGAGAACCTATCCTGAACTTACACT
Qy	101 ArgThrArg 103	Qy	21 ValGlnAsnValProTyroPheArgAspLysGlyAspSerAsnSerAlaGlyTrpMet
Db	993 CGTACACCT 1001	Db	761 AGGCCTTATGCTACTCCAGAACGGTTACTCTGACAAGTTGAATGGATG
RESULT 4	41 ValGlnAsnValProTyroPheArgAspLysGlyAspSerAsnSerAlaGlyTrpMet	820	80
AF020342	821 GTCAGATGTCCATACTTCAGGATAAGGGATTCAGTCAGCTGGATCGAACGTCAGTCAGCTGGATCGAACG	880	

QY 61 AsnSerIleArgHisAsnLeuSerIleSerLeuHisSerArgPheMetArgIleGlnAsnGluGly 80
 Db 81 AACTCCATCCCTCACATCCTCTTCATCTCGATTCAAGAAATCCGGTAACTGACTATCACT 795
 QY 81 AlaGlyLysSerSerTrpValLeuAsnProAspAlaLysProGlyMetAsnProArg 100
 Db 941 GCGGAAAGCTCGTGTTATAATCCAGTCAAAGCCGAAGAACCG 1000
 QY 101 ArgThrArg 103
 Db 1001 CGTACACGT 1009

RESULT 5
 AF032112 AF032112 3035 bp mRNA linear INV 19-NOV-1997
 LOCUS Caenorhabditis elegans DAF-16 (daf-16) mRNA, complete cds.
 DEFINITION AF032112
 ACCESSION AF032112.1 GI:2623942
 VERSION 1
 KEYWORDS Caenorhabditis elegans
 SOURCE Caenorhabditis elegans
 Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida;
 Rhabditoidae; Rhabditidae; Peloderaina; Caenorhabditis.
 REFERENCE 1. (bases 1 to 3035)
 AUTHORS Lin, K., Dorman, J.B., Rodan, A. and Kenyon, C.
 TITLE daf-16: An HNF-3/forkhead family member that can function to double
 the life-span of *Ceenorhabditis elegans*
 JOURNAL Science 278 (5341), 1319-1322 (1997)
 MEDLINE 98028757
 PUBMED 9360933
 REFERENCE 2. (bases 1 to 3035)
 AUTHORS Lin, K., Dorman, J.B., Rodan, A. and Kenyon, C.
 TITLE Direct Submission (30-OCT-1997) Biochemistry and Biophysics, UCSF, 513
 Submitted (30-OCT-1997) Biochemistry and Biophysics, UCSF, 513
 Panassus Ave., San Francisco, CA 94143-0554, USA
 JOURNAL
 FEATURES Location/Qualifiers
 source
 1. .3035
 /organism="Caenorhabditis elegans"
 /mol_type="mRNA"
 /strain="Bristol N2"
 /db_xref="Baxton:6239"
 /chromosome="I"
 gene
 1. .3035
 /gene="daf-16"
 /note="defective dauer formation; suppresses age phenotype
 of age-1, daf-2"
 334. .1866
 /note="HNF-3/forkhead homolog"
 /codon_start=1
 /product="DAF-16"
 /protein_id="AAC7803.1"
 /protein_id="GL:2623943"
 /translation="MMEMLVDOGTDASSAATSTSSVSRFGADTMNTPPDDVMMDDM
 APLPRURCNWPMRRIQLLEPPINNSPLIEQLEPDADLGSNEQQGQLGGASSNGST
 AMPLHTPDGSNTHOTSPPSDFMSSEDDTYSKKTTRRAWGNNSYAELITTAAMS
 DEKRULIAQYEMMVNVYPRDKEDNSNSAIGWNSIRNLNLSHFRMRONEGAKKS
 SWMVIDNDAKGRNPRTREKNTIETTQLERSKRKRRPALKMSLHSTLN
 GNSIASISSIONLSDLYDDSMQAFDNVPSSPRTOQSNSLPGSSSRVPAIGSDSY
 DDLEFSWWSGEVSVPALPSDIDVDRDMDRIDAATHIIGVOVKOESRPIKTEFLAPPY
 HELNSVRSQSCAQNPILRNIVPSINPKMPDLAGTNYQNGITDPLNLISTNSSSLDP
 G1QSGCGIVAAQTVASSALPDLNLTLPDPLMTMDVALLRHLSQGGQHJHF
 DL"

ORIGIN
 Alignment Scores:
 Pred. No.: 1.65e-51
 Score: 549.00
 Percent Similarity: 99.03%
 Best Local Similarity: 98.92%
 Query Match: DB:
 Length: 3035
 Matches: 102
 Conservative: 0
 Mismatches: 1
 Indels: 0
 Gaps: 0

misc_feature
 1
 /gene="fktf-1"
 /note="splice leader acceptor site; determined by binding
 of 5' RACE anchor primer with *Caenorhabditis elegans* SL1
 splice leader sequence"
 211. .2436
 /gene="fktf-1"
 /note="fktf-1a;
 /note="DNA-binding; alternatively spliced"
 /codon_start=1

	/db_xref="FLYBASE; FBgn0038197"
21	ThrAlaIleMetAlaSerProGluLysArgLeuThrIleAlaGlnValtryGluTrpMet 40
715	CATGCCATTGAGTCGCCACGACAACATTGACACTGATCAGATTCACTGAGTTGGATG 774
41	ValGlnAsnValProtYrRheAaspLysGlyAspSerAsnSerSerAlaGlyTrpLys 60
775	GTCAGAAATGTGCCATTTCAGGACAGGGCATTCGAATAGCAGTCGGTGGAG 834
61	AsnSerIleArgHisAsnLeuSerLeuIleSerArgPheMetArgIleGlnAsnGluGly 80
835	AACCTCCATAGTCACATTCAGTCGCAACCGCTTATGAGGTCTCAAAGGGCC 894
81	AlaGlyLysSerSerTPTPvalIleAsnProAspAlaLysProGlyMetAsnProArg 100
895	ACCGCCAAAGTCATCCTGGTCATGTCACCGGGCAAGCTGTGCCTGGCC 954
101	Arg 101
955	CGC 957
	ORIGIN
	ESTIMATE 8
	AY095452 mRNA 394 bp linear INV 18-MAR-2002
	LOCUS Drosophila melanogaster LDI1919 full insert cDNA.
	VERSION AV095452 GI:19528330
	FEATURES F1U_CDNA.
	ORGANISM Drosophila melanogaster (fruit fly)
	REFERENCE Drosophila: Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha; Drosophilidae; Drosophila; Drosophila.
	AUTHORS Steapleton, M., Brokstein, P., Hong, L., Agbayani, A., Carlson, J., Champe, M., Chavez, C., Dorrestein, V., Dresne, D., Farfan, D., Frisse, B., George, R., Gonzalez, M., Guarin, H., Kronmiller, B., Li, P., Liao, G., Miranda, A., Mungall, C. J., Munro, J. J., Paclob, J. J., Paragas, V., Park, S., Patel, S., Phouanenavong, S., Wan, K., Yu, C., Lewis, S. E., Rubin, G. M., and Celniker, S.
	TITLE Direct Submission
	JOURNAL Submitted (12-MAR-2002) Berkeley Drosophila Genome Project, Lawrence Berkeley National Laboratory, One Cyclotron Road, Berkeley, CA 94720, USA
	SEQUENCE submitted by: Berkeley Drosophila Genome Project
	Berkeley National Laboratory
	COMMENT This clone was sequenced as part of a high-throughput process to sequence clones from Drosophila Gene Collection 1 (Rubin et al., Science 2000). The sequence has been subjected to integrity checks for sequence accuracy, presence of a polyA tail and contiguity within 100 kb in the genome. Thus we believe the sequence to reflect accurately this particular cDNA clone. However, there are artifacts associated with the generation of cDNA clones that may have not been detected in our initial analyses such as internal priming, priming from contaminating genomic DNA, retained introns due to reverse transcription of unspliced precursor RNAs, and reverse transcriptase errors that result in single base changes. For further information about this sequence, including its location and relationship to other sequences, please visit our Web site (http://fruitfly.berkeley.edu) or send email to cDNA@fruitfly.berkeley.edu .
	FEATURES source
	ORGANISM "Drosophila melanogaster"
	/mol type="mRNA"
	/st="ain" r="1"; cn="bw" sp=""
	/db_xref="Taxon:7227"
	/map="18BA8-8BA9"
	1..3934
	LOCATIONS
	1..3934
	QUALIFIERS
	1. /organism="Drosophila melanogaster"
	/mol type="mRNA"
	/st="ain" r="1"; cn="bw" sp=""
	/db_xref="Taxon:7227"
	/map="18BA8-8BA9"
	1. /gene="foxo"
	/note="Alignment with genomic scaffold AE003703"
	gene
	/note="foxo"
	/note="Alignment with genomic scaffold AE003703"
	source
	ESTIMATE 8
	AY095452 mRNA 394 bp linear (1-3934)
	LOCUS Drosophila melanogaster AY095452 x AY095452 (1-3934)
	FEATURES alignment_scores:
	Pred. No.: 2.74e-40 Length: 3934
	Score: 450.00 Matches: 78
	Percent Similarity: 93.07% Conservative: 16
	Best Local Similarity: 77.22% Mismatches: 7
	Query Match: 81.08% Indels: 0
	DB: 3 Gaps: 0
	ORGANISM Drosophila melanogaster
	REFERENCE AY095452
	AUGAGAACATCTGGGTGGCAATGGATGGGAATCTATGGATCTCATACG 1099
	1. lysylthiethylarginylasparaginylalanyltyrosylalanine
	DB 1040 AAAGAACATCTGGGTGGCAATGGATGGGAATCTATGGATCTCATACG 1099
	21 thralailemetalaserasproglutylsargleuthienolealaGlnValtryGluTrpMet 40
	QY 1.1100 CATGCCATTGGATGGCACACAGATTGAGCTGAGTCAGATTACAGTGATGAT 1159
	41 valGlnasnyalprotYrPheArgAspLysGlyAspSerAsnSerAlaGlyTrpLys 60
	QY 1.1160 GTCCAAGATGGCCATTTCAGGAAAGGGGATGTCGAATAGCGCTGGATGAAAG 1219
	61 AsnSerIleArgHisAsnLeuSerIleHisSerArgPheMetArgIleGlnAsnGluGly 80
	DB 1220 AACTCCATAACCTCACAACTCTGCTGCAACACGGCTCAAACAGGGC 1279
	81 AlaGlyLysSerSerTrpTrpValIleAsnProAspAlaLysProGlyMetAsnProArg 100
	DB 1280 ACCGGGAAAGTCATCTGGATGCTCAACCCGGAAAGCTGCTGCGC 1339
	101 Arg 101
	Db 1340 CGC 1342
	RESULT 9
	AK092357 LOCUS AK092357 2685 bp mRNA linear
	DEFINITION mRNA clone OCBBF2016841, highly similar to FORKHEAD PROTEIN OSA.
	ACCESSION AK092357
	VERSION AK092357_1 GI:21750933
	KEYWORDS oligo capping; fts (full insert sequence).
	SOURCE Homo sapiens (human)
	ORGANISM Homo sapiens
	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Butelostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
	REFERENCE 1 Nishi,T., Ota,T., Nakagawa,S., Senoh,A., Mizuguchi,H., Inagaki,H., Sugiyama,T., Irie,R., Okuda,T., Sato,S., Wakamatsu,A., Ishii,S., Yamamoto,J., Isono,Y., Kawai,H., Saito,K., Nishikawa,T., Kimura,K., Yamashita,H., Maezuo,K., Nakamura,Y., Sekine,M.,
	AUTHORS

Kikuchi,H., Kanda,K., Wagatsuma,M., Murakawa,K., Kanehori,K., Sugiyama,A., Oshima,A., Kawakami,B., Suzuki,Y., Sugano,S., Nagahara,K., Matsuo,Y., Nagai,K. and Isogai,T.	SOURCE ORGANISM	Homo sapiens (human) Homo sapiens			
NEDO human cDNA sequencing project Unpublished	REFERENCE AUTHORS	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homiridae; Homo.			
2 (bases 1 to 2685)	JOURNAL TITLE AUTHORS JOURNAL TITLE COMMENT	Ota,T., Nakagawa,S., Senoh,A., Mizuguchi,H., Inagaki,H., Sugiyama,T., Irie,R., Otuki,T., Sato,H., Wakamatsu,A., Ishii,S., Yamamoto,J., Isono,Y., Kawai-Hio,Y., Saito,K., Nishikawa,T., Kazusa-Kamatari, Yamashita,H., Matsuo,K., Nakamura,Y., Sekine,M., Kikuchi,H., Kanda,K., Wagatsuma,M., Murakawa,K., Kanehori,K., Takahashi-Fujii,N., Oshima,A., Sugiyama,A., Kawakami,B., Suzuki,Y., Sugano,S., Nagahori,K., Matsuo,Y., Nagai,K. and Isogai,T.			
Submitted (04-JUN-2002) Takao Isogai, FLJ Project (HRI Team); 2-6-7 Kazusa-Kamatari, Kisarazu, Chiba 292-0812, Japan (E-mail: genomics@rii.co.jp, Tel: 81-438-52-3975, Fax: 81-438-52-3986) NEDO human cDNA sequencing project supported by Ministry of Economy, Trade and Industry of Japan; cDNA full insert sequencing: Research Association for Biotechnology (RAB); cDNA library construction: Helix Research Institute (HRI); (supported by Japan Key Technology Center etc.); 5' - & 3' -end one pass sequencing: RAB, HRI, and Biotechnology Center, National Institute of Technology and Evaluation; clone selection for full insert sequencing: HRI and RAB; annotation: HRI and RAB.	REFERENCE AUTHORS TITLE JOURNAL TITLE COMMENT	NEDO human cDNA sequencing project supported by Ministry of Economy, Trade and Industry of Japan; cDNA full insert sequencing: Research Association for Biotechnology (RAB); cDNA library construction: Helix Research Institute (HRI); (supported by Japan Key Technology Center etc.); 5' - & 3' -end one pass sequencing: RAB, HRI, and Biotechnology Center, National Institute of Technology and Evaluation; clone selection for full insert sequencing: HRI and RAB; annotation: HRI and RAB.			
FEATURES source	FEATURES source	1. "2870" /organism="Homo sapiens" /mol_type="mRNA" /db_xref="taxon:9606" /clone="BRTIA3007532" /tissue_type="thalamus" /clone.Lib="BRTIA3" /dev_stage="fetal" /note="cloning vector: pME18SFL3"			
ORIGIN	ORIGIN	1. "2870" /organism="Homo sapiens" /mol_type="mRNA" /db_xref="taxon:9606" /clone="BRTIA3007532" /tissue_type="thalamus" /clone.Lib="BRTIA3" /note="cloning vector: pME18SFL3"			
Alignment Scores:	Alignment Scores:				
Prod. No. : Score: Percent Similarity: Best Local Similarity: Query Match: DB:	Pred. No.: Score: Percent Similarity: Best Local Similarity: Query Match: DB:	2.08e-39 440.50 90.20% 77.45% 79.37% 9	2.24e-39 440.50 90.20% 77.45% 79.37% 9	Length: Matches: Conservative: Mismatches: Indels: Gaps:	2870 79 13 9 1 1
US-09-844-353A-54 (1-103) x AK093357 (1-2685)	US-09-844-353A-54 (1-103) x AK1228861 (1-2870)				
Qy 1 LysLysThrThrThrArgArgAspAlaArgPGLysAsnMetSerTyrAlaGluLeuIleThr 20 Db 146 AGGAATGGCTGCGGCCCTACTCAGGATAAGGGCACGCCAACCTTCATGCACCTGATC 205	Qy 21 ThrAlaIleMetAlaSerProGluLysArgLeuThrLeuAlaGlnValTyrGluTrpMet 40 Db 206 CGCCCATGGAGCTCCGGACAAACGGCTACTCTCCAGATCTACGAGTGGATG 265	Qy 41 ValGlnAsnValProTyrPheArgAspLysGlyAspSerAsnSerSerAlaGluLeuIleThr 20 Db 266 GTGGGTTGGGTGGCTACTCAGGATAAGGGCACGCCAACCTTCATGCACCTGATC 325	Qy 41 ValGlnAsnValProTyrPheArgAspLysGlyAspSerAsnSerSerAlaGluLeuIleThr 20 Db 341 AGGAAATGGCTGCGGCCGGACGCCAACCTTCATGCACCTGATC 400		
Qy 61 AsnSerIleArgIhsasnLeuSerLeuHisSerArgPheMetArgIleGlnAsnGluGly 80 Db 326 AACTCCATCGGCCAACCTGTACTGATAGTCATGGGTCAGATGAGCA 385	Qy 21 ThrAlaIleMetAlaSerProGluLysArgLeuThrLeuAlaGlnValTyrGluTrpMet 40 Db 401 CGGGCATGGAGCTCCGGACAAACGGCTACTCTCCAGATCTACGAGTGGATG 460	Qy 81 AlaGlyLysSerIleTrpValLeuSerLeuProAsp---AlaLysProGlyMetAsnPro 99 Db 386 ACTGGCAGAGCTTGGATCATACCCTGATAGTCATGGGAGAGGGAAAAGCCC 445	Qy 41 ValGlnAsnValProTyrPheArgAspLysGlyAspSerAsnSerSerAlaGluLeuIleThr 20 Db 461 GTGGTGGCGGCCCTACTTCAGGATAAGGGCACAGCTTGCCGCTGGAG 520		
Qy 100 ArgArg 101 Db 446 CGGGGG 451	Qy 61 AsnSerIleArgIhsAsnLeuSerLeuHisSerArgPheMetArgIleGlnAsnGluGly 80 Db 521 AACATCCATCGGCCAACCTGTACTCTCCAGATCTACGAGTGGATGAGGA 580	Qy 81 AlaGlyLysSerIleTrpValLeuSerLeuProAsp---AlaLysProGlyMetAsnPro 99 Db 581 ACTGGCAGAGCTTGGATCATACCCTGATGGGAGAGGGAAAAGCCC 640	RESULT 10 AK1228861 LOCUS DEFINITION Homo sapiens cDNA FLJ16486 f1, clone BRTHA007532, moderately similar to Forkhead box protein O3A. ACCESSION AK1228861 VERSION GI:34528058 KEYWORDS oligo capping; f1s (full insert sequence).		

Pred. No.:	2.51e-39	Length:	3183				
Score:	440.50	Matches:	79				
Percent Similarity:	90.20%	Conservative:	13				
Best Local Similarity:	77.45%	Mismatches:	9				
Query Match:	79.37%	Indels:	1				
DB:	9	Gaps:	1				
US-09-844-353A-54 (1-103) x AF032886 (1-3183)							
Qy	1 LysLysThrThrArgArgAsnAlaLysGlyAsnMetSerTyrAlaGluLeuLeuThr 20						
Db	1366 AGAAATGGTCGCGGGAGGCCGAACTGGATCGGCACCTGATCACCC 1425						
Qy	21 ThrAlaLeuMetAlaSerProGluLysArgLeuThrLeuAlaGlnValThrGluTrpMet 40						
Db	1426 CGGCCATGGAGCTCCTGGAAACCTCACTCTGCCAGATCTACGAGTGGATG 1485						
Qy	41 ValGlnAsnValProTyrrhepheAspGlyAspSerAlaGlyTrpLys 60						
Db	1486 GTCGTTGCGCCCTACTTCAAGATAAGGGCAAGCAACGCTCTGGCTGGGANG 1545						
Qy	61 AsnSerIearGlyIasnLeuSerLeuSerIearGlyIasnGluGly 80						
Db	1546 AACCCACCGGGACAACTGTACTGATAGCTGATGCTGATGCGGGTCAAATGAGGGA 1605						
Qy	81 AlaGlyLysSerSerIearGlyIasnProAsp---AlaLysProGlyMetAsnPro 99						
Db	1606 ACTGGCAAGGTCTGGATCATCACCTGATGAGCTAACCTGATCAAGATA 1665						
Qy	100 ArgArg 101						
Db	1666 CGGGGG 1671						
RESULT 13							
LOCUS	BC020227	3300 bp	mRNA	linear	PRI 06-OCT-2003		
DEFINITION	Homo sapiens forkhead box O3A, mRNA (CDNA clone MGC:31925						
IMAGE:	4111403)						
BC020227							
BC020227.1	complete cds.						
GI:	18044232						
MGC.							
Homo sapiens (human)							
ORGANISM							
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;							
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.							
1 (bases 1 to 3300),							
DEFINITION							
strausberg,R.L.,							
Klausner,R.D.,							
Alirschul,S.P.,							
Hopkins,R.F.,							
Ditachenko,L.,							
Stapleton,M.,							
Scheiner,T.E.,							
Carninci,P.,							
Abramson,R.D.,							
McFarland,M.C.,							
Worley,K.C.,							
Villalon,D.K.,							
Fahy,J.,							
Bonifard,G.G.,							
Dictionary,Y.S.,							
Schnarch,A.,							
TITLE							
JOURNAL	Proc. Natl. Acad. Sci. U.S.A. 99 (26), 16899-16903 (2002)						
MEDLINE							
PUBMED	22388257						
REFERENCE	12477932						
AUTHORS	2 (bases 1 to 3300)						
TITLE	Direct Submission						
JOURNAL	Submitted (19-DEC-2001) National Institutes of Health, Mammalian						
REMARK	Gene Collection (MGC), Cancer Genomics Office, National Cancer Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2530, USA						
COMMENT	NIH-MGC Project URL: http://mgc.nci.nih.gov						
	Contact: MGC help desk Email: cgabbs-r@mail.nih.gov						
	Tissue Procurement: ATCC						
	CDNA Library Preparation: Rubin Laboratory						
	CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)						
	DNA Sequencing by: Genome Sequence Centre, BC Cancer Agency, Vancouver, BC, Canada						
	info@bccgc.bc.ca						
	Steven Jones, Jennifer Asano, Ian Bosdet, Yaron Butterfield, Susanna Chan, Reaghan Chiu, Chris Fjell, Erin Garland, Ran Guin, Letricia Hsiao, Martin Krzywinski, Reta Kutsche, Oliver Lee, Soo Sen Lee, Victor Liang, Carrie Mathewson, Candice McLeavy, Steven Ness, Pavan Pandoh, Anna-Luisa Prabhu, Parvaneh Saeedi, Jacqueline Schein, Duane Smallus, Michael Smith, Lorraine Spence, Jeff Stott, Michael Thorne, Miranada Tsai, Natasha van den Bosch, Jill Vardy, George Yang, Scott Zuyderduyn, Marco Marra.						
	Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at: http://image.llnl.gov						
	Series: IRAL Plate: 40 Row: c Column: 5						
	This clone was selected for full length sequencing because it passed the following selection criteria: matched mRNA gi: 4503738.						
	Location/Qualifiers						
	1 .. 3300						
	/organism="Homo sapiens"						
	/mol type="mRNA"						
	/db_xref="taxon:9606"						
	/clone="MGC:31925 IMAGE:4111403"						
	/tissue type="muscle, rhabdomyosarcoma"						
	/clone_id="NMG_17"						
	/lab_host="DH10B-R"						
	/note="vector: pOTB7"						
	gene						
	/note="synonyms: FKHL11, AF6g21, FKHLIP2"						
	/db_xref="Taxon:9606"						
	/clone_id="NMG_17"						
	/lab_host="DH10B-R"						
	/note="vector: pOTB7"						
	gene						
	/note="synonyms: FOXO3A"						
	/db_xref="Taxon:9606"						
	/clone_id="NMG_17"						
	/lab_host="DH10B-R"						
	/note="vector: pOTB7"						
	gene						
	/note="synonyms: FOXO3A"						
	/db_xref="Taxon:9606"						
	/clone_id="NMG_17"						
	/lab_host="DH10B-R"						
	/note="vector: pOTB7"						
	gene						
	/note="synonyms: FKHL11, AF6g21, FKHLIP2"						
	/db_xref="Taxon:9606"						
	/clone_id="NMG_17"						
	/lab_host="DH10B-R"						
	/note="vector: pOTB7"						
	gene						
	/note="synonyms: FOXO3A"						
	/db_xref="Taxon:9606"						
	/clone_id="NMG_17"						
	/lab_host="DH10B-R"						
	/note="vector: pOTB7"						
	gene						
	/note="synonyms: FOXO3A"						
	/db_xref="Taxon:9606"						
	/clone_id="NMG_17"						
	/lab_host="DH10B-R"						
	/note="vector: pOTB7"						
	gene						
	/note="synonyms: FOXO3A"						
	/db_xref="Taxon:9606"						
	/clone_id="NMG_17"						
	/lab_host="DH10B-R"						
	/note="vector: pOTB7"						
	gene						
	/note="synonyms: FOXO3A"						
	/db_xref="Taxon:9606"						
	/clone_id="NMG_17"						
	/lab_host="DH10B-R"						
	/note="vector: pOTB7"						
	gene						
	/note="synonyms: FOXO3A"						
	/db_xref="Taxon:9606"						
	/clone_id="NMG_17"						
	/lab_host="DH10B-R"						
	/note="vector: pOTB7"						
	gene						
	/note="synonyms: FOXO3A"						
	/db_xref="Taxon:9606"						
	/clone_id="NMG_17"						
	/lab_host="DH10B-R"						
	/note="vector: pOTB7"						
	gene						
	/note="synonyms: FOXO3A"						
	/db_xref="Taxon:9606"						
	/clone_id="NMG_17"						
	/lab_host="DH10B-R"						
	/note="vector: pOTB7"						
	gene						
	/note="synonyms: FOXO3A"						
	/db_xref="Taxon:9606"						
	/clone_id="NMG_17"						
	/lab_host="DH10B-R"						
	/note="vector: pOTB7"						
	gene						
	/note="synonyms: FOXO3A"						
	/db_xref="Taxon:9606"						
	/clone_id="NMG_17"						
	/lab_host="DH10B-R"						
	/note="vector: pOTB7"						
	gene						
	/note="synonyms: FOXO3A"						
	/db_xref="Taxon:9606"						
	/clone_id="NMG_17"						
	/lab_host="DH10B-R"						
	/note="vector: pOTB7"						
	gene						
	/note="synonyms: FOXO3A"						
	/db_xref="Taxon:9606"						
	/clone_id="NMG_17"						
	/lab_host="DH10B-R"						
	/note="vector: pOTB7"						
	gene						
	/note="synonyms: FOXO3A"						
	/db_xref="Taxon:9606"						
	/clone_id="NMG_17"						
	/lab_host="DH10B-R"						
	/note="vector: pOTB7"						
	gene						
	/note="synonyms: FOXO3A"						
	/db_xref="Taxon:9606"						
	/clone_id="NMG_17"						
	/lab_host="DH10B-R"						
	/note="vector: pOTB7"						
	gene						
	/note="synonyms: FOXO3A"						
	/db_xref="Taxon:9606"						
	/clone_id="NMG_17"						
	/lab_host="DH10B-R"						
	/note="vector: pOTB7"						
	gene						
	/note="synonyms: FOXO3A"						
	/db_xref="Taxon:9606"						
	/clone_id="NMG_17"						
	/lab_host="DH10B-R"						
	/note="vector: pOTB7"						
	gene						
	/note="synonyms: FOXO3A"						
	/db_xref="Taxon:9606"						
	/clone_id="NMG_17"						
	/lab_host="DH10B-R"						
	/note="vector: pOTB7"						
	gene						
	/note="synonyms: FOXO3A"						
	/db_xref="Taxon:9606"						
	/clone_id="NMG_17"						
	/lab_host="DH10B-R"						
	/note="vector: pOTB7"						
	gene						
	/note="synonyms: FOXO3A"						
	/db_xref="Taxon:9606"						
	/clone_id="NMG_17"						
	/lab_host="DH10B-R"						
	/note="vector: pOTB7"						
	gene						
	/note="synonyms: FOXO3A"						
	/db_xref="Taxon:9606"						
	/clone_id="NMG_17"						
	/lab_host="DH10B-R"						
	/note="vector: pOTB7"						
	gene						
	/note="synonyms: FOXO3A"						
	/db_xref="Taxon:9606"						
	/clone_id="NMG_17"						
	/lab_host="DH10B-R"						
	/note="vector: pOTB7"						
	gene						
	/note="synonyms: FOXO3A"						
	/db_xref="Taxon:9606"						
	/clone_id="NMG_17"						
	/lab_host="DH10B-R"						
	/note="vector: pOTB7"						
	gene						
	/note="synonyms: FOXO3A"						
	/db_xref="Taxon:9606"						
	/clone_id="NMG_17"						
	/lab_host="DH10B-R"						
	/note="vector: pOTB7"						
	gene						
	/note="synonyms: FOXO3A"						
	/db_xref="Taxon:9606"						
	/clone_id="NMG_17"						
	/lab_host="DH10B-R"						
	/note="vector: pOTB7"						
	gene						
	/note="synonyms: FOXO3A"						
	/db_xref="Taxon:9606"						
	/clone_id="NMG_17"						
	/lab_host="DH10B-R"						
	/note="vector: pOTB7"						
	gene						
	/note="synonyms: FOXO3A"						
	/db_xref="Taxon:9606"						
	/clone_id="NMG_17"						
	/lab_host="DH10B-R"						
	/note="vector: pOTB7"						
	gene						
	/note="synonyms: FOXO3A"						
	/db_xref="Taxon:960						

This Page Blank (uspto)

DRAFT

Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on: July 29, 2004, 09:24:12 ; Search time 341 Seconds

(without alignments)

1 KKTTRRNANGNMSTYELIT.....SSWWVWINPDAKPGMNP RRTR 103

Title: US-09-844-353A-54

Perfect score: 555

Sequence: 1 KKTTRRNANGNMSTYELIT.....SSWWVWINPDAKPGMNP RRTR

Scoring table: BLOSUM62

Xgapop	10.0	Xgapext	0.5
Ygapop	10.0	Ygapext	0.5
Fgapop	6.0	Fgapext	7.0
Delop	6.0	Delext	7.0

Searched: 3373863 seqs, 2124099041 residues

Total number of hits satisfying chosen parameters:

6747726

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Command line parameters:

```
-MODEL:frame_p2n.model -DEV=xlp
-Q /cgn2/1/OSPT spool p=/USO944353/runat 23072004 084913_28228/app_query.fasta_1.263
-DB=N Geneseq_29cpan04 -QFMT=fastap -SUFFIX=rng -MINNATC=0 -LOOP2CL=1
-LCOPENTR=0 -UNITS=bits -START=1 -END=1 -MATRIX=blossom62 -TRANS=human40.cdi
-LIST=45 -DOCALL=200 -THR SCORR=pct -THR MAX=100 -THR MIN=0 -ALIGNN=15
-MODE=LOCAL -OUTfmt=pcl -NORM=ext -HFAPOSTL=500 -MINLEN=0 -MAXLEN=2000000000
-USER=USO944353 @CGN_1_1_885 @runat 23072004 084913_28228 -NCPU=6 -ICPP=3
-NO_MMW -LARGEQUERY -NEGPBLCK=100 -LONGLOG
-DEV TIMEOUT=120 -WARN TIMEOUT=10 -THREADS=1 -XGAPPOP=10 -XGAPEXT=0.5 -FGAPPOP=6
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELEXT=7
```

RESULT 1

ID AAA62298 standard; cdna; 3499 BP.

XX XX AAA62298;

AC AC

DT DT 11-JAN-2001 (first entry)

XX XX Caenorhabditis elegans daf-16 CDNA #1.

DE DE

XX XX Caenorhabditis elegans; daf-16; daf-2; age-1; daf-18;

KW KW insulin signalling pathway; insulin receptor; PI 3-kinase; PKB kinase;

KW KW AKT kinase; PREN lipid phosphatase; antidiabetic; anorectic; obesity;

KW KW diabetes; impaired glucose tolerance; transgenic animal; ss.

XX XX OS OS Caenorhabditis elegans.

XX XX Location/Qualifiers

FT FT 325. -185

/tag= a

/product= "DAF-16"

/transl_except= (pos:679. .681,aa:Pro)

/transl_except= (pos:682. .684,aa:Ser)

/transl_except= (pos:685. .687,aa:Asp)

/note= "the codon at position 1207 to 1208 has an apparent 1 nucleotide deletion, which alters the reading frame."

1	555	100.0	3499	3	AAA62298	Aaa62298 Caenorhab
2	422.5	77.0	5723	6	ABY94143	Abv94143 Breast ca
3	422.5	76.1	2067	7	ACD19349	Acd19349 cdna enc
4	410.5	74.0	1073	4	AAS0244	Aas0244 Fusion pr
5	410.5	74.0	3387	4	AAI60219	Aai60219 Human pol
6	410.5	74.0	3394	4	AAI58433	Aai58433 Human pol
7	410.5	74.0	3394	8	ADB48404	Adb48404 Novel hum
8	406.5	73.2	3171	6	ABS73211	Abs73211 DNA encod

PD 08-JUN-2000.

Pred. No. is the number of results predicted by chance to have a a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query	Match Length	DB ID	Description
1	555	100.0	3499	3	AAA62298 Caenorhab
2	422.5	77.0	5723	6	ABY94143 Breast ca
3	422.5	76.1	2067	7	ACD19349 cdna enc
4	410.5	74.0	1073	4	AAS0244 Fusion pr
5	410.5	74.0	3387	4	AAI60219 Human pol
6	410.5	74.0	3394	4	AAI58433 Human pol
7	410.5	74.0	3394	8	ADB48404 Novel hum
8	406.5	73.2	3171	6	ABS73211 DNA encod

Pred. No. is the number of results predicted by chance to have a a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query	Match Length	DB ID	Description
1	555	100.0	3499	3	AAA62298 Caenorhab
2	422.5	77.0	5723	6	ABY94143 Breast ca
3	422.5	76.1	2067	7	ACD19349 cdna enc
4	410.5	74.0	1073	4	AAS0244 Fusion pr
5	410.5	74.0	3387	4	AAI60219 Human pol
6	410.5	74.0	3394	4	AAI58433 Human pol
7	410.5	74.0	3394	8	ADB48404 Novel hum
8	406.5	73.2	3171	6	ABS73211 DNA encod

XX ABV94143;
 XX AC
 XX DT 08-JAN-2003 (first entry)
 XX DE Breast carcinoma related nucleotide sequence SEQ ID NO:134.
 XX Human; breast carcinoma; cancer; tumour; cytosstatic; anti-tumour; gene;
 XX KW ss.
 XX KW
 XX OS Homo sapiens.
 XX PN WO200246467-A2.
 XX PD 13-JUN-2002.
 XX PP 07-DEC-2001; 2001WO-IB002811.
 XX PR 08-DEC-2000; 2000US-0254090P.
 XX PR 07-DEC-2001; 2001US-0007926.
 XX PA (IPSO-) IPSOGEN.
 XX PI Bertucci F, Houlgate R, Birnbaum D, Nguyen C, Viens P, Fert V;
 XX PI
 XX DR WO/2002-619023/66.
 XX Novel polynucleotide library useful in molecular characterization of a carcinoma, comprising a pool of polynucleotide sequences or its subsequences which are either underexpressed or overexpressed in tumor cells.
 XX Claim 1: Page 198-200; 401pp; English.
 XX The present invention describes a polynucleotide library (I) useful in the molecular characterisation of a carcinoma, comprising a pool of polynucleotides or its subsequences which are either underexpressed or overexpressed in tumour cells, and correspond to any of the polynucleotide sequences chosen from the 468 sequences given in ABV94010 to ABV94477. Also described: (1) a polynucleotide array (II) useful for the prognosis or diagnosis of tumour, comprising (I); and (2) detecting (M1) differentially expressed polynucleotide sequences which are correlated with a cancer, involves obtaining a polynucleotide sample from a patient, and reacting the polynucleotide sample obtained with a probe immobilised on a solid support, where the probe comprises any combination of the polynucleotide sequences of (I) or its expression products encoded by polynucleotide sequences of (II), and detecting the reaction product. (I) have cytostatic activities and can be used as anti-tumour agents. (I) is useful in molecular characterisation of a carcinoma. (I) and (II) are useful for large-scale molecular characterisation of tumour, in differentiating a normal cell from a cancer cell, detecting a hormone sensitive tumour cell, differentiating a tumour with lymph nodes from a tumour without lymph nodes, differentiating antibiotic-sensitive tumours from antibiotic-insensitive tumours, and classifying good and poor prognosis primary breast tumours. (I) is useful for large-scale molecular characterisation of breast cancer that help in prediction, prognosis and cancer treatment, and for detecting differentially expressed genes that correlate with a cancer
 XX Sequence 3499 BP; 990 A; 805 C; 565 G; 1139 T; 0 U; 0 Other;
 XX Alignment Scores:
 XX Preid. No.: 9.25E-63 Length: 3499
 XX Score: 555.00 Matches: 103
 XX Percent Similarity: 100.00% Conservative: 0
 XX Best Local Similarity: 100.00% Mismatches: 0
 XX Query Match: 100.00% Indels: 0
 XX DB: 3 Gaps: 0
 US-09-844-353A-54 (1-103) x AAAG62298 (1-13499)
 Qy 1 LysLysthrThrArgAsnAlaTrpGlyAsnMetSerTyraGluLeuIleThr 20
 Db 727 AAAAGAACGACCAAGCGAACGGTGGAAATATGTCTATGTCAGTTACT 786
 Qy 21 ThralAlaMetAlaSerProGluLysArgLeuThrLeuAlaGlnValTyrGluTrpMet 40
 Db 787 ACGCCATTATGCCTAGTCAGCAAGTTAACCTTGACAAGTTCAATGATG 846
 Qy 41 ValGlnAsnValProTrpPheArgAsnLysGlyAspSerAsnSerAlaGlyTrpLys 60
 Db 847 GTCCAGATGATGCCATCTTCGGGATAAGGATTCAGTTAGCTGATGAAAG 906
 Qy 61 AsnSerIleArgHisIleSerLeuHisSerArgPheMetArgIleGlnAsnGluGly 80
 Db 907 AACTGATGCCGTCACATCTGNTCTCATTCGTAATCAGATGAAATCAGATGAAAG 966
 Qy 81 AlaGlyLysSerSerTrpTrpValleAsnProAspAlaLysProGlyMetAsnProArg 100
 Db 967 GCGGAAGAGCTGCTGGTGGTTATTAATCCAGATGAAAGCAGGAATGATCCCGG 1026
 Qy 101 ArgThrArg 103
 Db 1027 CGTACACGT 1035
 US-09-844-353A-54 (1-103) x ABV94143 (1-5723)
 Qy 1 LysLysThrThrArgArgAsnAlaTrpGlyAsnMetSerTyraGluLeuIleThr 20
 RESULT 2
 ABV94143 standard; cDNA; 5723 BP.
 ID ABV94143

DE Fusion protein GST-AfxDBD; DNA sequence.
 XX DNA binding domain; Afx; transcription factor; GST-AfxDBD; ds;
 KW human fork head transcription factor; Afx response element; human;
 KW diabetes; drug target; insulin; insulin receptor signalling pathway.
 XX
 OS Homo sapiens.
 OS Synthetic.
 OS Chimeric.
 XX
 Key 1 . 1073
 CDS /*tag= a
 /product= "GST-AfxDBD fusion protein"
 /transl_except= (pos:82 . 84, aa:Thr)
 /transl_except= (pos:1046 . 1047, aa:Ser)
 /note= "This codon contains an apparent 1 nucleotide
 deletion which alters the reading frame"
 XX WO200114544-A1.
 XX PD 01-MAR-2001.
 XX PP 22-AUG-2000; 2000WO-SE001603.
 XX PR 26-AUG-1999; 99SE-000003009.
 PR 31-AUG-1999; 99US-0151167P.
 PR 25-NOV-1999; 99SE-00004269.
 XX PA (PHAA) PHARMACIA & UPJOHN AB.
 XX Climent-Johansson I, Dahlman-Wright K, Lake S, Wasserman W,
 FI DR WPI; 2001-218446/22.
 DR P-PSDB; AAU00214.
 XX PT New Afx response element with a nucleotide sequence comprising a DNA
 PT binding site for the human fork head transcription factor Afx, useful in
 PT screening for genes or in bioinformatic analysis of the human genome.
 XX Example 2; Fig 7; 34pp; English.
 XX The sequence represents the coding sequence of the GST-AfxDBD fusion
 CC protein encoded by expression vector pGEX-AfxDBD. Human fork head
 CC transcription factor, Afx, was expressed and the protein used to find a
 CC response element comprising an 8 base pair (bp) nucleotide sequence. The
 CC nucleotide sequence comprises AACATCT, the Afx response element, and is
 CC useful in bioinformatic analysis e. g. of the human genome. Employing the
 CC Afx response element is also useful for screening for genes that may be
 CC used as diabetes drug targets. This can provide a subset of genes
 CC transcriptionally responsive to insulin and may lead to development of
 CC assays that facilitate the analysis of genes interacting with the insulin
 CC receptor pathway. The genes found in such screening may also be employed
 CC in further screening methods for compounds which modify the insulin
 CC receptor signalling pathway.
 XX Sequence 1073 BP; 300 A; 235 C; 268 G; 270 T; 0 U; 0 Other;
 SQ Alignment Scores:
 Fred. No.: 2.56e-44 Length: 1073
 Score: 41.0-50 Matches: 73
 Percent Similarity: 86.27% Conservative: 15
 Best Local Similarity: 71.57% Mismatches: 13
 Query Match: 73.96% Indels: 1
 DB: 4 Gaps: 1
 US-09-844-353A-54 (1-103) x AAS00244 (1-1073)
 Qy 1 LysLysThrThrArgArgAsnAlaTrpGlyAsnMetSerTyrAlaGluLeuIleThr 20
 ::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||:
 Db 715 CGGAAGGGCTCCGGGAATCAGTCATATGGAACTCATGAGC 774
 Qy 21 ThralAlleMetAlaSerProGluLysArgLeuThrLeuAlaGlnValtryglutTrpMet 40

Db 775 CAGGCCATTGAAAGGCCCGGAAAGCGACTGACACTTGCCAGATCTACGAGTGGATG 834
 Qy 41 ValGlnAsnValProlotyPheArgAspLysGlyAspSerAsnSerSerIleGlnTrpLys 60
 Db 835 GTCCGTACTGTTACCTTACTCTCAAGGACACGGTACAGCAACAGCTCAGAGGGAG 894
 Qy 61 AsnSerIleArgHsAsnIleSerLeuHisSerIleAsnGluGly 80
 Db 895 AACTGATGGCACACCTGGCAGCAGTTCAGCAGGTTCAACAGGGCC 954
 Qy 81 AlaGlyLysserSerIrrTrpValleAsnProAsp--AlaLysProGlyMetAsnPro 99
 Db 955 ACCGGAAAAGCTTGGTGAATGCTGGAGGCAAGGC 1014
 Qy 100 ArgArg 101
 Db 1015 CGCGC 1020

RESULT 5
 AA160219 ID AA160219 standard; cDNA; 3387 BP.
 XX
 XX AC AA160219;
 XX DT 22-OCT-2001 (first entry)
 XX DE Human polynucleotide SEQ ID NO 4208.
 XX KW Human; nootropic; immunosuppressant; cytostatic; gene therapy; cancer;
 KW peripheral nervous system; neuropathy; central nervous system; CNS;
 KW Alzheimer's; Parkinson's disease; Huntington's disease; haemostatic;
 KW amytrophic lateral sclerosis; Shy-Drager Syndrome; chemotactic;
 KW chemokinetic; thrombolytic; drug screening; arthritis; inflammation;
 KW leukaemia; ss.
 XX OS Homo sapiens.
 XX PN WO200153312-A1.
 XX PD 26-JUL-2001.
 XX PF 26-DEC-2000; 2000WO-US034263.
 XX PR 23-DEC-1999; 99US-00471275.
 PR 21-JAN-2000; 2000US-00488725.
 PR 25-APR-2000; 2000US-00552317.
 PR 20-JUN-2000; 2000US-00598042.
 PR 19-JUL-2000; 2000US-00620312.
 PR 03-AUG-2000; 2000US-00653450.
 PR 14-SEP-2000; 2000US-00662191.
 PR 19-OCT-2000; 2000US-00693036.
 PR 29-NOV-2000; 2000US-00727344.
 XX PA (HYSEQ INC.)
 XX Tang YT, Liu C, Asundi V, Chen R, Ma Y, Qian XB, Ren F, Wang D, Zhao QA,
 PI Wang J, Wang Z, Wehrum T, Xu C, Xie AJ, Yang Y, Zhang J, Zhang J, Zhao QA,
 PI Zhou P, Goodrich R,
 XX DR WPI; 2001-442253/47.
 DR P-PSDB; AAU01063.
 XX PT Novel nucleic acids and polypeptides, useful for treating disorders such
 as central nervous system injuries.
 XX PS Claim 1; SBQ ID NO 4208; 10078pp; English.
 XX The invention relates to human nucleic acids (AAI57798-AAI61369) and the
 CC encoded polypeptides (AAI38642-AM2213) with nootropic,
 CC immunosuppressant and cytostatic activity. The polynucleotides are useful
 CC in gene therapy. A composition containing a polypeptide or polynucleotide
 CC of the invention may be used to treat diseases of the peripheral nervous

CC system, such as peripheral nervous injuries, peripheral neuropathy and CC localised neuropathies and central nervous system diseases, such as CC Alzheimer's, Parkinson's disease, Huntington's disease, amyotrophic lateral sclerosis, and Shy-Drager Syndrome. Other uses include the CC utilisation of the activities such as: Immune system suppression, CC Activin/inhibin activity, chemokinetic/chemokinetic activity, haemostatic and thrombolytic activity, cancer diagnosis and therapy, drug screening, CC assays for receptor activity, arthritis and inflammation, leukaemia and CC C.N.S disorders. Note: The sequence data for this patient did not form CC part of the printed specification

XX Sequence 3387 BP; 785 A; 891 C; 1000 G; 711 T; 0 U; 0 Other;

Alignment Scores:
Pred. No.: 1 32e-43 Length: 3387
Score: 41.050 Matches: 73
Percent Similarity: 86.27% Conservative: 15
Best Local Similarity: 71.57% Mismatches: 13
Query Match: 73.96% Indels: 1
DB: 4 Gaps: 1

US-09-844-353A-54 (1-103) x AAI60219 (1-3387)

QY 1 LysLysThrThrArgArgAsnAlaTrpGlyAsnMetSerTyrAlaGluIleLeuThr 20
Db 690 CGAAAGGGAGGCTCCGCCGGAATGCCCTGGAAATCAGTCATATGCCGAACTCATGG 749
QY 21 ThrAlaIleMetAlaSerProGluLysArgLeuThrLeuAlaGlnValtryGluTrpMet 40
Db 750 CAGGCCATTGAAAGGCCCGGAGAACGGACTGACACTTGCCCAAGAATCTACCAAGTGGATG 809
QY 41 ValGlnAsnValProtryPheArgAspLysGlyAspSerAsnSerSerAlaGlyTrpLys 60
Db 810 GICCGTACTGTACCTACTCTCAAGGCAAGGGTACAGCTAGCATGAGTGAAG 869
QY 61 AsnSerIleArgHisAsnLeuSerLeuHisSerArgPheMetArgIleGlnAsnGluGly 80
Db 870 AACTCGATCGGACAACTGTCCCTGACAAGTTCAAGGTTCAACAGCGAGGCC 929
QY 81 AlaGlyLysSerSerSerTrpTrpValIleAsnProAsp--AlaLysProGlyMetAsnPro 99
Db 930 ACCGGCAAAGCTCTTGTGATGCTGAAACCTCTGAGGGAAAGGGCAAAGGCC 989

CC PR 21-JAN-2000; 20000US-00488725.
CC PR 25 APR 2000; 20000US-00552317.
CC PR 20-JUN-2000; 20000US-00598042.
CC PR 19-JUL-2000; 20000US-00620312.
CC PR 03-AUG-2000; 20000US-00653450.
CC PR 14-SEP-2000; 20000US-00662191.
CC PR 19-OCT-2000; 20000US-00693036.
CC PR 29-NOV-2000; 20000US-00727344.
XX PA (HYSE-) HYSEQ INC.

XX PR Tang YT, Liu C, Asundi V, Chen R, Ma Y, Qian XB, Ren F, Wang D;
PI Wang J, Wang Z, Weinman T, Xu C, Xue AJ, Yang Y, Zhang J, Zhao QA;
PI Zhou P, Goodrich R, Drmanac RT;
XX DR WPI; 2001-442253/47.
XX DR P-PSDB; AAM39277.
PT Novel nucleic acids and polypeptides, useful for treating disorders such
as central nervous system injuries.
XX PT
PS Claim 1; SEQ ID NO 636; 10078pp; English.
XX

CC The invention relates to human nucleic acids (AAI57798-AAI61369) and the
CC encoded polypeptides (AAI38642-AAM2213) with nootropic,
CC immunosuppressant and cyrostatic activity. The polynucleotides are useful
CC in gene therapy. A composition containing a polypeptide or polynucleotide
CC of the invention may be used to treat diseases of the peripheral nervous
CC system, such as peripheral nervous injuries, peripheral neuropathy and
CC localised neuropathies and central nervous system diseases, such as
CC Alzheimer's, Parkinson's disease, Huntington's disease, amyotrophic
CC lateral sclerosis, and Shy-Drager Syndrome. Other uses include the
CC utilisation of the activities such as: Immune system suppression,
CC Activin/inhibit activity, chemokinetic/chemokinetic activity, haemostatic
CC and thrombolytic activity, cancer diagnosis and therapy, drug screening,
CC assays for receptor activity, arthritis and inflammation, leukaemias and
CC C.N.S disorders. Note: The sequence data for this patient did not form
part of the printed specification

XX SQ Sequence 3394 BP; 790 A; 893 C; 999 G; 712 T; 0 U; 0 Other;

Alignment Scores:
Pred. No.: 1 1.32e-43 Length: 3394
Score: 41.050 Matches: 73
Percent Similarity: 86.27% Conservative: 15
Best Local Similarity: 71.57% Mismatches: 13
Query Match: 73.96% Indels: 1
DB: 4 Gaps: 1

US-09-844-353A-54 (1-103) x AAI58433 (1-3394)

QY 1 LysLysThrThrArgArgAsnAlaTrpGlyAsnMetSerTyrAlaGlyValtryGluIleThr 20
Db 690 CGGAAGGGAGGCTCCGCCGGAATGCCCTGGAAATCAGTCATATGCCGAACTCATGG 749
QY 21 ThrAlaIleMetAlaSerProGluLysArgLeuThrLeuAlaGlnValtryGluTrpMet 40
Db 750 CAGGCCATTGAAAGGCCCGGAGAACGGACTGACACTTGCCCAAGAATCTACCAAGTGGATG 809
QY 41 ValGlnAsnValProtryPheArgAspLysGlyAspSerAsnSerSerAlaGlyTrpLys 60
Db 810 GICCGTACTGTACCTACTCTCAAGGCAAGGGTACAGCTAGCATGAGTGAAG 869
QY 61 AsnSerIleArgHisAsnLeuSerLeuHisSerArgPheMetArgIleGlnAsnGluGly 80
Db 870 AACTCGATCCGCCAACCTGTCAGTTCAAGGTTCAACAGCGAGGCC 929
QY 81 AlaGlyLysSerSerTrpTrpValIleAsnProAsp--AlaLysProGlyMetAsnPro 99
Db 930 ACCGGAAAAGCTCTTGTGATGCTGAAACCTCTGAGGGAAAGGGCAAAGGCC 989

RESULT 6
AAI58433
ID AAI58433 standard; cDNA; 3394 BP.
XX AC AAI58433;
XX DT 22-Oct-2001 (first entry)
XX Human Polynucleotide SEQ ID NO 636.

DE DB 750 CAGGCCATTGAAAGGCCCGGAGAACGGACTGACACTTGCCCAAGAATCTACCAAGTGGATG 809
QY 21 ThrAlaIleMetAlaSerProGluLysArgLeuThrLeuAlaGlnValtryGluTrpMet 40
Db 750 CAGGCCATTGAAAGGCCCGGAGAACGGACTGACACTTGCCCAAGAATCTACCAAGTGGATG 809
QY 61 AsnSerIleArgHisAsnLeuSerLeuHisSerArgPheMetArgIleGlnAsnGluGly 80
Db 870 AACTCGATCCGCCAACCTGTCAGTTCAAGGTTCAACAGCGAGGCC 929
QY 81 AlaGlyLysSerSerTrpTrpValIleAsnProAsp--AlaLysProGlyMetAsnPro 99
Db 930 ACCGGAAAAGCTCTTGTGATGCTGAAACCTCTGAGGGAAAGGGCAAAGGCC 989

XX Homo sapiens.
XX WO200153312-A1.
XX PD 26-JUL-2001.
XX PF 26-DEC-2000; 2000WO-US034263.
XX PR 23-DEC-1999; 99US-00471275.
QY 100 ArgArg 101

CC synovial sarcoma. The method is also useful for treating viral infections. This represents the DNA sequence of a chromosome aberration

SQ Sequence 3171 BP; 742 A; 855 C; 909 G; 665 T; 0 U; 0 Other;

Alignment Scores:

Pred. No. :	4.07e-43	Length:	3171
Score:	406.50	Matches:	72
Percent Similarity:	85.29%	Conservative:	15
Best Local Similarity:	70.59%	Mismatches:	14
Query Match:	73.24%	Indels:	1
DB:	6	Gaps:	1

US-09-844-353A-54 (1-103) × ABS73211 (1-3171)

QY	1 LysLysThrThrArgArgAsnAlaTrpGlyAsnMetSerTyAlaGluLeuIleThr 20
Db	444 CGGAAGGGGGCCCGGGAAATGCCCTGGGAATTCAGTCATAATGCAGAACATCAGC 503
QY	21 ThralleMetLeaserProGluLysArgLeuThrLeuAlaGlnValtryGluTrpMet 40
Db	504 CAGGCCATTGAAAGGCCCGGGAGGCACTGACATGCGAGATTCAGGTGGATG 563
QY	41 ValGlnAsnValProTyPheArgAspLysGlyAspSerAsnSerSerAlaGlyTriPlys 60
Db	564 GTCCGTACTGTACCTACTTCAGGAAACGGTGACAGAACAGCTCACAGGATGGAG 623
QY	61 AspSerIleArgHisAsnLeuSerLeuHisSerArgPheMetArgIleGlnAsnGluGly 80
Db	624 AACTCGATCGGCCAACCTGTCCCTGCAACAGGTTAACAGGTTAACAGGGCC 683
QY	81 AlaGlyLysSerSerTrpTrpValLeasnProasp---AlaLysProGlyMetAsnPro 99
Db	684 ACCGGCAAAAGCTTGGGGATGCTGAACCTTGAGGGCAAGGCC 743
QY	100 ArgArg 101
Db	744 CGCGC 749

RESULT 9
ABS73212
ID ABS73212 standard; DNA; 3171 BP.
XX AC ABS73212;
XX DT 04-DEC-2002 (First entry)
XX DE DNA encoding human translocation (X; 11) (q13; q23) protein #2.
XX KW Chromosome aberration; oncogenic fusion protein; cancer; proliferative disease; cellular protein isoform; heat shock protein 90; KW KW; rheumatoid arthritis; cancer; haematopoietic disorder; T cell lymphoma; B cell lymphoma; chronic myeloid leukaemia; CML; KW acute myeloid leukaemia; AML; NHL; solid tumour; papillary thyroid carcinoma; ELL; APL; CMML; KW rhabdomyosarcoma; synovial sarcoma; Ewing's sarcoma; melanoma; liposarcoma; KW Homo sapiens.
XX OS Homo sapiens.
XX PN WO200269900-A2.
XX PD 12-SEP-2002.
XX PF 01-MAR-2002; 2002WO-US006518.
XX PR 01-MAR-2001; 2001US-0272751P.
XX PA (CONF-) CONFORMA THERAPEUTICS CORP.
XX PI Fritz LC, Burrows FJ;
XX PN WPI; 2002-698710/75.
DR

DR p-PSDB; ABG95043.

XX PT Treating genetically-defined disease associated with chromosomal aberrations yielding oncogenic fusion proteins, e.g. cell proliferative diseases, involves administering an inhibitor of heat shock protein 90.

XX PS Disclosure: Page 157; 38pp; English.

CC The invention describes a method of treating genetically-defined disease associated with chromosomal aberrations yielding oncogenic fusion proteins (I), treating cancerous cells containing (I) in a heterogeneous cell population, treating proliferative diseases associated with mutant protein or cellular protein isoforms (II) dependent on heat shock protein (HSP)-90, or selectively treating cells expressing (II) involving administering HSP90-inhibitor. The method is useful for treating oncogenic fusion protein, treating cancerous cells containing fusion protein in heterogeneous cell population, treating proliferative disease (e.g. rheumatoid arthritis or cancer) associated with mutant protein or cellular protein isoform dependent on heat shock protein (HSP)-90 (e.g. p53), or selectively treating cells expressing mutant protein or cellular protein isoform in a patient heterozygous for (II). The method is useful for treating a disease e.g. haematopoietic disorder such as T or B cell lymphoma, chronic myeloid leukaemia (CML), APL, ALL, AML and CMML, or a disease characterised by a solid tumour such as papillary thyroid carcinoma, Ewing's sarcoma, melanoma, liposarcoma, rhabdomyosarcoma and synovial sarcoma. The method is also useful for treating viral infections. This represents the DNA sequence of a chromosome aberration

QY	Sequence 3171 BP; 742 A; 855 C; 909 G; 665 T; 0 U; 0 Other;
Db	US-09-844-353A-54 (1-103) × ABS73212 (1-3171)
QY	1 LysLysThrThrArgArgAsnAlaTrpGlyAsnMetSerTyAlaGluLeuIleThr 20
Db	444 CGGAAGGGGGCCCGGGAAATGCCCTGGGAATTCAGTCATAATGCAGAACATCAGC 503
QY	21 ThralleMetLeaserProGluLysArgLeuThrLeuAlaGlnValtryGluTrpMet 40
Db	504 CAGGCCATTGAAAGGCCCGGGAGGCACTGACATGCGAGATTCAGGTGGATG 563
QY	41 ValGlnAsnValProTyPheArgAspLysGlyAspSerAsnSerSerAlaGlyTriPlys 60
Db	564 GTCCGTACTGTACCTACTTCAGGAAACGGTGACAGAACAGCTCACAGGATGGAG 623
QY	61 AspSerIleArgHisAsnLeuSerLeuHisSerArgPheMetArgIleGlnAsnGluGly 80
Db	624 AACTCGATCGGCCAACCTGTCCCTGCAACAGGTTAACAGGTTAACAGGGCC 683
QY	81 AlaGlyLysSerSerTrpTrpValLeasnProasp---AlaLysProGlyMetAsnPro 99
Db	684 ACCGGCAAAAGCTTGGGGATGCTGAACCTTGAGGGCAAGGCC 743
QY	100 ArgArg 101
Db	744 CGCGC 749

QY	1 LysLysThrThrArgArgAsnAlaTrpGlyAsnMetSerTyAlaGluLeuIleThr 20
Db	444 CGGAAGGGGGCCCGGGAAATGCCCTGGGAATTCAGTCATAATGCAGAACATCAGC 503
QY	21 ThralleMetLeaserProGluLysArgLeuThrLeuAlaGlnValtryGluTrpMet 40
Db	504 CAGGCCATTGAAAGGCCCGGGAGGCACTGACATGCGAGATTCAGGTGGATG 563
QY	41 ValGlnAsnValProTyPheArgAspLysGlyAspSerAsnSerSerAlaGlyTriPlys 60
Db	564 GTCCGTACTGTACCTACTTCAGGAAACGGTGACAGAACAGCTCACAGGATGGAG 623
QY	61 AsnSerIleArgHisAsnLeuSerLeuHisSerArgPheMetArgIleGlnAsnGluGly 80
Db	624 AACTCGATCGGCCAACCTGTCCCTGCAACAGGTTAACAGGTTAACAGGGCC 683
QY	81 AlaLysSerSerTrpTrpValLeasnProasp---AlaLysProGlyMetAsnPro 99
Db	684 ACCGGCAAAAGCTTGGGGATGCTGAACCTTGAGGGCAAGGCC 743
QY	100 ArgArg 101
Db	744 CGCGC 749

RESULT 10	AAS85053
ID	AAS85053 standard; cDNA; 5833 BP.
XX	AC AAS85053 ;
XX	DT 13-FEB-2002 (first entry)

DE	DNA encoding novel human diagnostic protein #20857.	Db	960 GGTAAATGGTCCCTACTTCAGGATAAGGTAAGCTGGCAACAGCTCGGCCTGGAA 1019
XX	Human; chromosome mapping; gene mapping; gene therapy; forensic; ss.	QY	60 sAsnSerIleArgHisAsnLeuSerLeuIleSerArgPheMetArg-TleGlnnsGluG 80
KW	food supplement; medical imaging; diagnostic; genetic disorder; ss.	Db	1020 GAATTCATTCGTTATAATCTGTCCTACAGGAAGTCATTCCTGAGATGANG 1079
XX	Homo sapiens.	QY	80 lYAlaGlyIysSerSerTrpValleAsnProAsp---AlaLysProGlyMetAsn 99
OS		Db	1080 GAACTGGAAAAGTCTGGATGCTAAATCCAGGGTGGCAAGAGGGAAATCIC 1139
XX	WO200175067-A2.	QY	99 roAFArG 101
PD	11-OCT-2001.	Db	1140 CTAGGAGA 1147
XX	30-MAR-2001; 2001WO-US008631.	RESULT 11	
XX	31-MAR-2000; 2000US-00540217.	AAA67299	
PR	23-AUG-2000; 2000US-00649167.	ID	AAA67299 standard; cDNA; 2704 BP.
XX	(HYSEQ) HYSEQ INC.	XX	
PA	Drmanac RT, Liu C, Tang YT;	AC	AAA67299;
XX	DR ; 2001-639362/73.	XX	XX 11-JAN-2001 (First entry)
DR	P-PSDB; AEG20866.	DE	Caenorhabditis elegans daf-16 cDNA #2.
XX	New isolated polynucleotide and encoded polypeptides, useful in diagnostics, forensics, gene mapping, identification of mutations responsible for genetic disorders or other traits and to assess biodiversity.	XX	XX Caenorhabditis elegans; daf-16; daf-2; age-1; daf-18; PKB kinase; insulin signalling pathway; insulin receptor; PI 3-kinase; AKT kinase; PTEN lipid phosphatase; antidiabetic; anorectic; obesity; diabetes; impaired glucose tolerance; transgenic animal; ss.
PT	Claim 1; SEQ ID NO 20857; 103pp; English.	XX	XX
PR	CC The invention relates to isolated polynucleotide (I) and polypeptide (II) sequences, (I) is useful as hybridisation probes, polymerase chain reaction (PCR) primers, oligomers, and for chromosome and gene mapping, and in recombinant production of (II). The polynucleotides are also used in diagnostics as expressed sequence tags for identifying expressed genes. (I) is useful in gene therapy techniques to restore normal activity of (II), or to treat disease states involving (II). (II) is useful for generating antibodies against it, detecting or quantitating a polypeptide in tissue, as molecular weight markers and as a food supplement. (II) and its binding partners are useful in medical imaging of sites expressing (II). (I) and (II) are useful for treating disorders involving aberrant protein expression or biological activity. The polypeptide and polynucleotide sequences have applications in diagnostics, forensics, gene mapping, identification of mutations responsible for genetic disorders or other traits to assess biodiversity and to produce other types of data and products dependent on DNA and amino acid sequences. AAS64197-AAS9564 represent novel human diagnostic coding sequences of the invention. Note: The sequence data for this patent did not appear in the printed specification, but was obtained in electronic format directly from WIPO at ftp.wipo.int/pub/published_pct_sequences	OS	OS Caenorhabditis elegans.
PS	XX Sequence 5833 BP; 1482 A; 1383 C; 1331 G; 1637 T; 0 U; 0 Other;	FH	Location/Qualifiers
SQ	XX	FT	14 .1543
	XX	FT	/*tag= a
	XX	FT	/product= "DAF-16"
	XX	PN	WO200033068-A1.
	XX	PD	08-JUN-2000.
	XX	XX	XX
	XX	PF	02-DEC-1999; 99WO-US028529.
	XX	PR	03-DEC-1998; 98US-00205658.
	XX	XX	XX (GEHO) GEN HOSPITAL CORP.
	XX	PI	Ruvikun G, Ogg S;
	XX	DR	WPI; 2000-423022/36.
	XX	DR	P-PSDB; AAB06071.
	XX	XX	XX Diagnosing and treating obesity and impaired glucose tolerance using modulators of daf-18 expression and/or activity.
	XX	PT	PT
	XX	XX	XX Disclosure; Fig 13B; 402pp; English.
	XX	XX	The present cDNA sequence from Caenorhabditis elegans corresponds to a differentially spliced daf-16 transcript which encodes an isoform of DAF-16. DAF-16 interacts with the C. elegans insulin signalling pathway. A number of C. elegans genes have been identified as homologues of genes in the mammalian insulin signalling pathway. The C. elegans age-1 gene encodes a homologue of the mammalian PI 3-kinase whilst daf-2 encodes a homologue of the mammalian insulin receptor. The C. elegans PI 3-kinase and AKT kinase act downstream of daf-2 and age-1, just as their mammalian homologues act downstream of insulin signalling. Other daf genes have also been implicated in the C. elegans insulin signalling pathway. The C. elegans PTEN lipid phosphatase homologue, DAF-18, has been found to act upstream of AKT in the pathway. This discovery has enabled mammalian PTEN action to be mapped to the insulin signalling pathway. Compounds that inhibit the expression and/or activity of polypeptides encoded by these genes may be administered to patients to treat or prevent disorders such as obesity and impaired glucose tolerance.
QY	Alignment Scores:	Length:	5833
	Bred. No.:	Matches:	73
	Score:	Conservative:	17
	Percent Similarity:	Mismatches:	1
	Local Similarity:	Indels:	3
	Query Match:	Gaps:	1
DB:	5		
QY	US-09-844-353A-54 (1-103) x AAS85053 (1-5833)		
QY	1 LysLysThrThrThrArgAsnAlaTrpGlyAsnMetSerTyraGluLeuIleThr 20		
DB:	840 AAAGCACGCTCGNCCCCCGAACCTGTCTTACGCCGACCTCATACC 899		
QY	21 ThrAlaIleMetLaserProGluLysArgLeuThrLeuAlaGlnVal-TyrGluTrpMe 40		
DB:	900 AACGCCATGAGCTCGCGAGAACGGCTCACGGTGTGCOAGATCTAACGAGTGAT 959		
QY	40 tValGlnAsnValProtYrrPheArgAspLysGlyAspSerAsnSerAlaGlyTrpLy 60		

XX	SQ Sequence 2704 BP; 709 A; 669 C; 452 G; 874 T; 0 U; 0 Other;	XX	PT New human AFX-zeta polypeptide, useful for identifying substances that antagonize or agonize the activity of AFX-zeta protein for treating a disease associated with AFX-zeta protein, e.g. diabetes, atherosclerosis or dyslipidemia.
Alignment Scores:		PT	antagonize or agonize the activity of AFX-zeta protein for treating a disease associated with AFX-zeta protein, e.g. diabetes, atherosclerosis or dyslipidemia.
Pred. No. :	1.68e-39	PT	
Score:	378.50	PT	
Percent Similarity:	75.9%	PT	
Best Local Similarity:	62.04%	PT	
Query Match:	68.20%	PS	Claim 5; Page 35; 48pp; English.
DB:	3	XX	
US-09-844-353A-54 (1-103) x AAA62299 (1-2704)		CC	The present sequence is the coding sequence for AFX-zeta, a novel splice variant of human transcription factor AFX. AFX-zeta was cloned from human liver and heart libraries. The encoded protein has a deletion of 55 amino acids as compared with AFX-alpha protein (see ABP72186), including the first 16 amino acids of the forkhead domain of the AFX-alpha protein.
Qy 1 LyslysthrThr-----ThrArgArgAspAlaTrpGlyAsnMetSerIyr 15		CC	Reporter gene assays showed that AFX-zeta is a potent transcription activator with properties distinct from AFX-alpha. It is regulated by the insulin signalling pathway, binding insulin responsive sequence (IRS) fragments, and by an agent known to affect AMPK activity. It may be a distal effector of the insulin-signalling pathway. AFX-zeta mRNA shows strongest expression in liver, kidney and pancreas. The invention provides assay techniques for identifying human AFX-zeta agonists and antagonists. These can be used to treat and/or prevent insulin resistance, hyperglycaemia, hypoglycaemia, hepatic gluconeogenesis, hypercholesterolaemia, obesity, type 2 diabetes, impaired glucose tolerance, disorders related to syndrome X including hypertension, obesity, insulin resistance, coronary artery disease, glomerulonephritis, glomerulosclerosis, nephrotic syndrome, arthritis, Parkinson's disease, Alzheimer's disease, Huntington's disease, atherosclerosis, nephrosclerosis, and dyslipidaemia.
Db 398 AAGAACCCACCGATCAATTGGCAGAGAAACCGAACATGGATCTTAT 457		CC	
Qy 16 AlaGluLeuIleThrThrAlaLeuMetAlaSerProGluLysArgLeuThrLeuAlaGln 35		CC	
Db 458 TCGATATCATGGCAAAGGATTGGAATCGGCCAGACGGAACGCTTAACCTAAATGAG 517		CC	
Qy 36 ValTrpGluTrpMetValGlnAsnValProTyroPheArgAspLysGlyAspSerAsnSer 55		CC	
Db 518 ATTATCATGGTCCTCTGATAATAATTCCCTACTTGTGAAACGATCTAGTCCCGAGAG 577		CC	
Qy 56 SerAlaGlyTrpLysAsnSerIleArgHisAsnLeuSerLeuHisSerArgPheMetArg 75		CC	
Db 578 GCGCGGTTGGAAACTGATGTCATCTGTCATCTGTTGATGCA 637		CC	
Qy 76 IleGlnAsnGluGlyAlaGlyLysSerSerTrpTrpValIleAsnProAspAlaLysPro 95		CC	
Db 638 ATTGAAATGAAAGGCGGAAAGAGCTGTGGTTATTATCCAGATGCAAAGCCA 697		CC	
Qy 96 GlyMetAsnProArgArgThrArg 103		XX	Sequence 1353 BP; 295 A; 355 C; 355 G; 258 T; 0 U; 0 Other;
Db 698 GGATGAAATCCACGGCGTAACGCT 721		XX	
RESULT 12		XX	
ABZ58161	ID ABZ58161 standard, cDNA; 1353 BP.	XX	
AC ABZ58161;		XX	
DT 22-APR-2003 (first entry)		XX	
XX Human transcription factor AFX-zeta splice variant cDNA.		DE	
XX AFX-zeta; human; transcription factor; splice variant; antidiabetic; antiarteriosclerotic; antihypertensive; cytoprotective; neuroprotective; anticonvulsant; anticardiotonic; anorectic; hypotensive; nephrotropic; cardiotonic; gene therapy; gene; ss.		KW	
XX Homo sapiens.		OS	
XX Key Location/Qualifiers		XX	
PH 1.1353	1.	XX	
CDS /* tag= a		XX	
FT /product= "Human AFX-zeta"		XX	
FT WO2003002601-A2.		XX	
PN 09-JAN-2003.		XX	
XX 28-JUN-2002; 2002WO-EP007160.		PF	
XX 29-JUN-2001; 2001US-0302134P.		PR	
XX (NOV) NOVARTIS AG.		XX	
XX (NOV) NOVARTIS-ERFINDUNGEN VERW GES MBH.		XX	
PA Bowen BR, Yang Z, Whelan JP;		PI	
XX DR WPI; 2003-210242/20.		XX	
KW Drosophila; developmental biology; cell signalling; insecticide; pharmaceutical; gene; ss.		DR	

XX	Drosophila melanogaster.	1138	GCGCTTCCACGGATTTCAGCTTCTCCCTCGCCGGTTGCTGGTGGTCCGT 1197
OS	WO2001171042-A2.	60	-----
XX	WPI; 2001-656860/75.	60	-----
PN	P-PSDB; ABB61582.	60	-----
XX	New isolated nucleic acid detection reagent for detecting 1000 or more genes from Drosophila and for elucidating cell signaling and cell-cell interactions.	60	-----
PT	SEQ ID NO 11537; 21pp + Sequence Listing; English.	60	-----
CC	The invention relates to an isolated nucleic acid detection reagent capable of detecting 1000 or more genes from Drosophila. The invention is useful in developmental biology and in elucidating cell signalling and cell-cell interactions in higher eukaryotes for the development of insecticides, therapeutics and pharmaceutical drugs. The invention discloses genomic DNA sequences (AB16176-AB13051), expressed DNA sequences (AB101840-AB16175) and the encoded proteins (ABB5737-CCB272). The sequence data for this patent did not form part of the printed specification, but was obtained in electronic format directly from WIPO at fp.wipo.int/pub/published_pct_sequences	60	-----
XX	Sequence 2162 BP; 574 A; 602 C; 553 G; 433 T; 0 U; 0 Other;	60	-----
SQ	Alignment Scores: Pred. No.: 7.24e-34 Score: 335.00 Percent Similarity: 33.57% Best Local Similarity: 27.86% Query Match: 60.36% DB: 4	Length: 2162 Matches: 78 Conservative: 16 Mismatches: 7 Indels: 179 Gaps: 1	DNB05684 standard; cDNA; 25891 BP. XX ABLO5684; XX ABLO5684; XX DT 26-MAR-2002 (first entry) XX DE Drosophila melanogaster expressed polynucleotide SEQ ID NO 11534. XX KW Drosophila; developmental biology; cell signalling; insecticide; pharmaceutical; gene; ss. XX OS Drosophila melanogaster. XX PN WO200171042-A2. XX PD 27-SEP-2001. XX PF 23-MAR-2001; 2001WO-US009231. XX PR 23-MAR-2000; 2000US-0191637P. XX PR 11-JUL-2000; 2000US-00614150. XX PA (PEKE) PE CORP NY. XX PI Venter JC, Adams M, XX DR Wei; 2001-656860/75. XX DR P-PSDB; ABB61581. XX PT New isolated nucleic acid detection reagent for detecting 1000 or more genes from Drosophila and for elucidating cell signaling and cell-cell interactions. XX PS Claim 1; SEQ ID NO 11534; 21pp + Sequence Listing; English. XX CC The invention relates to an isolated nucleic acid detection reagent capable of detecting 1000 or more genes from Drosophila. The invention is useful in developmental biology and in elucidating cell signalling and cell-cell interactions in higher eukaryotes for the development of insecticides, therapeutics and pharmaceutical drugs. The invention
XX	1 LysLysthrThrArgAsnAlaIrrGlyAsnMetSerTyraGluLeuIleThr 20	60	-----
Db	778 ArgAAAGAACATCAGCTGCATGGCAATGGGGAAATCTATGGGATCTCATCAGC 837	60	-----
QY	21 ThrlAla11MetalaserProGluLysArgLeuThrAlaGlnValtryGluTrpMet 40	60	-----
Db	838 CATGCCATTGGATGCCAACGACAAGATGACACTGAGTCAGATTACAGTGATG 897	60	-----
QY	41 ValGlnAsnValProtryPheArgAspLysGlyAspSerAsnSerSerAlaGlyTrpLys 60	60	-----
Db	898 GTCCAGAATGTGCCATTTCAGGAZAGGGCATGATAGCAGTTACGAGTGGCAAG 957	60	-----
QY	958 AGTTTGGATAATGGCTTATCCTGGATCGTTTCGTTTACATGGAGC 1017	60	-----
Db	1018 ATACCACCTCTGGCTGTGATAAAAGACCAACGGCTGTGAAGAATTCGCTGCTCA 1077	60	-----
QY	CGAAAGAAATCGAAGACGACTCTTCACATTATTCGAAACTTATTCGCAAGAAGC 1137	60	-----
Db	1078 CGAAAGAAATCGAAGACGACTCTTCACATTATTCGCAAGAAGC 1137	60	-----
QY	-----	60	-----

discloses genomic DNA sequences (ABL16176-ABL30511), expressed DNA sequences (ABL01840-ABL16175) and the encoded proteins (ABB57737-ABB72072). The sequence data for this patent did not form part of the printed specification, but was obtained in electronic format directly from WIPO at [ftp://ftpt.wipo.int/pub/published_pct_sequences](http://ftpt.wipo.int/pub/published_pct_sequences)

Sequence 25891 BP; 7426 A; 5074 C; 5058 G; 8333 T; 0 U; 0 Other;

Alignment Scores:

Refred. No.:	1.21e-22	Length:	25891
Score:	262.00	Matches:	47
Percent Similarity:	91.94%	Conservative:	10
Deepest Local Similarity:	75.81%	Mismatches:	5
Query Match:	47.21%	Indels:	0
BB:	4	Gaps:	0

(S-09-844-353A-54 (1-103) × ABL05684 (1-25891)

1 LysLysThrThrArgArgAsnAlaTrpGlyAsnMetSerItyAlaGluLeuIleThr 20
11002 ArgAGAACTCATCGCTGCATGGAAATCATTCCATPGCGGATCATCAG 11061
21 ThrAlaIleMetAlaSerProGluLysArgLeuThrIleAlaClnValTyrGluTIPMet 40
11062 CATGCCATTGGATGGCAACGATGACAATGAGTCAGATTACGAGTGGATG 11121
41 ValGlnAsnValProTypheAspGlyAspGlyAspSerAsnSerSerAlaGlyTrp 60
11122 GTCCGAGATGTGCCATATTCAAGAACAGGGGATCAGATGGAGTGGCGATGGA 11181
61 AsnSer 62
11182 GTGAGT 11187

ESSENTIAL 15

AAD55840 standard; cDNA; 5123 BP.

AAD55840;

07-AUG-2003 (first entry)

Human nucleic acid associated protein (NAAP) -11 cDNA.

Human; nucleic acid associated protein; NAAP; stroke; AIDS; nootropic; cancer; atherosclerosis; neurological; epilepsy; Huntington's disease; developmental disorder; antiinflammatory; neuroprotective; thyromimetic; Cushing's syndrome; infection; gene therapy; cytoprotective; anticonvulsant; cerebroprotective; hypothyroidism; cell proliferative disorder; allergy; gene; ss.

Homo sapiens.

Key CDSS

Location/Qualifiers 179 . 1945 /*tag= a /product= "Human NAAP protein"

WO2003005618-A2.

23-JAN-2003.

X 1.0-JUL-2002; 2002WO-US021971.

X 12-JUL-2001; 2001US-0305089P.

R 12-JUL-2001; 2001US-0305104P.

R 13-JUL-2001; 2001US-0305345P.

R 13-JUL-2001; 2001US-0305390P.

R 19-JUL-2001; 2001US-030690P.

R 20-JUL-2001; 2001US-0306654P.

R 27-JUL-2001; 2001US-0308170P.

XX Swarnakar A, Richardson TW, Warren BA, Griffin JA, Tang YT; PI Yue H, Baugh MR, Emerling BM, Lal PG, Lu DAM, Forsythe IJ; PI Ramkumar J, Li JK, Becha SD, Duggan BM, Sanjivwala MM, Lee EA; PI Burford N, Elliott VS, Ison CH, Ding L, Barroso I, Tran B, Walia NK, Hafalia AJA, Nguyen DB, Lu Y; PI Arvizu CS; XX WPI; 2003-221732/21. DR P-PSDB; AAE37024.

XX New human nucleic acid associated proteins (NAAP), useful for diagnosing, treating and preventing diseases or conditions associated with the aberrant NAAP expression e.g. cancer, AIDS, atherosclerosis, epilepsy, or infections.

XX Claim 5; Page 239-241; 260pp; English.

XX The invention relates to human nucleic acid associated proteins (NAAP) and their corresponding nucleic acid sequences. The invention is useful in diagnosing, treating and preventing diseases or conditions associated with the decreased expression or overexpression of NAAP, such as cell proliferative (e.g. cancer, atherosclerosis), neurological (e.g. epilepsy, Huntington's disease, stroke), immune/inflammatory (e.g. AIDS, allergies) and developmental (e.g. Hypothyroidism, Cushing's syndrome) disorders, or infections. It is also useful in assessing the effects of exogenous compounds on the expression of nucleic acid and amino acid sequences of NAAP. The NAAP or its fragments are useful in screening compounds for which acts as their agonist or antagonist. The microarray is useful in monitoring or measuring protein-protein interactions, drug-target interactions, and gene expression profiles. NAAP DNA is used in gene therapy. The present sequence is human NAAP cDNA.

XX Sequence 5123 BP; 1383 A; 1103 C; 1115 G; 1522 T; 0 U; 0 Other; SQ Alignment Scores:

Pred. No.:	2.09e-18	Length:	5123
Score:	222.50	Matches:	45
Percent Similarity:	66.67%	Conservative:	19
Best Local Similarity:	46.88%	Mismatches:	21
Query Match:	40.09%	Indels:	11
DB:	7	Gaps:	4

US-09-844-353A-54 (1-103) × AAD55840 (1-5123)

QY 14 SerTyrAlaGluLeuIleThrThrAlaSerProGluLysArgLeuThrLeu 33
Db 422 AGTTATGCCAGCCTCAATTACATTGCAATAATAGCTACCCAAAAGAAAATGACTTA 481
QY 34 AlaGlnValTyrGluTrpMetValGlnAsnValProTypheAspLysGlyAspSer 53
Db 482 AGTCAAATTATCAGTGATGTTGATAACTCCATAATTAGAGGGTGGC----- 535
QY 54 AsnSerSerAlaGlyTriLysAsnSerIleArgHisAsnLeuSerLeuHisSerArgPhe 73
Db 536 -----AGTGTGTTGGAGAAATTCCATGACATAATCTGTCATTGACAAATGTTTC 586
QY 74 MetArgIle-----GlnAsnGluGlyAlaGlyLysSerSerTrpTrpValle----- 89
Db 587 CTTAAAGTGCCTGATCTAAGGTGACCTGCTACTGGCAATTAGACACC 646
QY 90 AsnPro-----AspAlaLysProGlyMetAsnProArgArgThrArg 103
Db 647 AATCCGAAAGAATGCGCTGATCTGGCAAAGAAAGGGCAAGCGA 694

Search completed: July 29, 2004, 10:40:16
Job time : 350 secs



Copyright (c) 1993 - 2004 Compugen Ltd.

protein - nucleic search, using frame_plus_p2n model

1 on: July 29, 2004, 10:27:35 ; Search time 81 Seconds
(without alignments)
705.678 Million cell updates/sec

file: US-09-844-353A-54
effect score: 555
sequence: 1 KKTITERRNAGNMYSAILIT.....SSWWVTPNDAKPGMNPRRTR 103

searching table: BLOSUM62

Xgapop	10.0	Xgapext	0.5
Ygapop	10.0	Ygapext	0.5
Egapop	6.0	Egapext	7.0
Delop	6.0	Delect	7.0

682709 seqs, 277475446 residues

total number of hits satisfying chosen parameters:
Maximum DB seq length: 0
Maximum DB seq length: 2000000000

post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

command line parameters:
IDEF-l-frame+ .p2n.model -DEV=x4P
/cgn2_1/USPROT/spool/p/IS09844353/runat_23072004_0849114_28266/app_query.fasta_1.263
-ISSUED_PATENTS_NA -QFORMAT=fasta -SUFFIX=rna -MINMATCH=0.1 -LOCFL=0
-QOPEXTI=0 -UNITS=6its -SPART=-1 -END=-1 -MATRIX=blosum62 -TRANS=blosum62 -ALIGN=15
-ST=45 -DOCALIGN=200 -THR SCORE=pct -THR MAX=100 -THR MIN=0 -MAXLEN=2000000000
-DEFLocal-OUTfmt=pto -NORM=ext -HEAPSIZE=2000000000 -MINLEN=1000000000
DEFLocal-OUTfmt=@CGN 1 1 128 @runat 23072004 0849114 28266 -NCPU=6 -ICPU=3
MMAP -LARGEQUERY -NEG SCORES=0 -WAIT DSPLOCK=100 -LONGLOG
TIMEOUT=1200 WARN TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6
-QOPEXTI=7 -YGAPOP=0.5 -DELOP=6 -DELECT=7

abbase : Issued_Patents_NA:*

1:	/cgn2_6/picodata/2/ina/5A_COMB.seq:*
2:	/cgn2_6/picodata/2/ina/5B_COMB.seq:*
3:	/cgn2_6/picodata/2/ina/6A_COMB.seq:*
4:	/cgn2_6/picodata/2/ina/6B_COMB.seq:*
5:	/cgn2_6/picodata/2/ina/PCTRIS_COMB.seq:*
6:	/cgn2_6/picodata/2/ina/backfiles1.seq:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

%
SUMMARIES

SUMM	Query	Match	Length	DB	ID	Description
1	555	100.0	3499	3	US-08-857-076-43	Sequence 43, Appl 1
2	410.5	74.0	1073	4	US-09-645-629-22	Sequence 22, Appl 1
3	410.5	74.0	3394	4	US-09-642-512D-14	Sequence 314, Appl 1
4	378.5	68.2	2704	3	US-08-857-076-44	Sequence 44, Appl 1
5	185.5	33.4	2830	1	US-07-882-292-1	Sequence 1, Appl 1
6	185.5	33.4	2830	2	US-08-331-44-1	Sequence 1, Appl 1
7	185.5	33.4	2830	5	PCT-US93-04102-1	Sequence 1, Appl 1
8	184.5	33.2	1860	2	US-08-331-64-3	Sequence 3, Appl 1
9	184.5	33.2	1860	5	PCT-US93-04102-3	Sequence 3, Appl 1
10	182.5	32.9	1659	3	US-09-083-351-3	Sequence 3, Appl 1
11	182.5	32.9	1659	11	US-09-083-351-3	Sequence 3, Appl 1
12	182.5	32.9	3946	3	US-09-083-351-1	Sequence 1, Appl 1

Copyright (c) 1993 - 2004 Compugen Ltd.

protein - nucleic search, using frame_plus_p2n model

1 on: July 29, 2004, 10:27:35 ; Search time 81 Seconds
(without alignments)
705.678 Million cell updates/sec

file: US-09-844-353A-54
effect score: 555
sequence: 1 KKTITERRNAGNMYSAILIT.....SSWWVTPNDAKPGMNPRRTR 103

searching table: BLOSUM62

Xgapop	10.0	Xgapext	0.5
Ygapop	10.0	Ygapext	0.5
Egapop	6.0	Egapext	7.0
Delop	6.0	Delext	7.0

searched: 682709 seqs, 277475446 residues

total number of hits satisfying chosen parameters:
Maximum DB seq length: 0
Maximum DB seq length: 2000000000

post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

command line parameters:
IDEF-l-frame+ .p2n.model -DEV=x4P
/cgn2_1/USPROT/spool/p/IS0984353/runat_23072004_0849114_28266/app_query.fasta_1.263
-ISSUED_PATENTS_NA -QFORMAT=fasta -SUFFIX=rna -MINMATCH=0.1 -LOCFL=0
-QOPEXTI=0 -UNITS=6its -SPART=-1 -END=-1 -MATRIX=blosum62 -TRANS=blosum62 -ALIGN=15
-ST=45 -DOCALIGN=200 -THR SCORE=pct -THR MAX=100 -THR MIN=0 -MAXLEN=2000000000
-DEFLD=LOCAL OUTFMT=pto -NORM=ext -HEAPSIZE=2000 -MINLEN=100
DEFLD=LOCAL OUTFMT=@CGN 1..1..128 @runat 23072004 0849114_28266 -NCPU=6 -1CPU=3
NAME -LARGEQUERY -NEG SCORES=0 -WAIT DSPLOCK=100 -LONGLOG
TIMEOUT=1200 WARN TIMEOUT=30 -THREADS=1 -XGAP=10 -XGAPEXT=0.5 -FGAPOP=6
-QOPEXTI=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELETE=7

abbase : Issued_Patents_NA.*
1: /cgn2_6/pctodata/2/ina/5A_COMB.seq: *
2: /cgn2_6/pctodata/2/ina/5B_COMB.seq: *
3: /cgn2_6/pctodata/2/ina/6A_COMB.seq: *
4: /cgn2_6/pctodata/2/ina/6B_COMB.seq: *
5: /cgn2_6/pctodata/2/ina/PCTRS_COMB.seq: *
6: /cgn2_6/pctodata/2/ina/backfile1.seq: *

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

%
SUMMARIES

abbase	Query	Match	Length	DB	ID	Description
1	555	100.0	3499	3	US-08-857-076-43	Sequence 43, Appl 1
2	410.5	74.0	1073	4	US-09-645-629-22	Sequence 22, Appl 1
3	410.5	74.0	3394	4	US-09-645-620-112D-14	Sequence 314, Appl 1
4	378.5	68.2	2704	3	US-08-857-076-44	Sequence 44, Appl 1
5	185.5	33.4	2830	1	US-07-882-292-1	Sequence 1, Appl 1
6	185.5	33.4	2830	2	US-08-331-44-1	Sequence 1, Appl 1
7	185.5	33.4	2830	5	PCT-US93-04102-1	Sequence 1, Appl 1
8	184.5	33.2	1860	2	US-08-331-64-3	Sequence 3, Appl 1
9	184.5	33.2	1860	5	PCT-US93-04102-3	Sequence 3, Appl 1
10	182.5	32.9	1659	3	US-09-083-351-3	Sequence 3, Appl 1
11	182.5	32.9	1659	11	US-09-083-351-3	Sequence 3, Appl 1
12	182.5	32.9	3946	3	US-09-083-351-1	Sequence 1, Appl 1

1 LysLysThrThrArgArgAsnAlaTyrGlyAsnMetSerTyrAlaGluLeuIleThr 20
 QY 727 AAAAGCAACAAGCACCAGACCGAACCTTGGAATAATGTCATAATGCTGACTTACT 786
 Db 787 ACAGCCATTATGGCTAGTCAGAAACGGTTAACTCTGCACAGTGTACGAATGATG 846

21 ThrlAlaIleMetAlaSerProGluLySArgLeuIleLeuAlaGlnValItyGluTrpMet 40
 QY 41 ValGlnAsnAlaProtYrPheArgAspLysGlyAspSerIasnSerSerAlaGlyTrpLys 60
 Db 847 GTCCAAATGTCATCTAGGGTAAGGGAGATTGAGACAGTTCAGTTGAG 906

61 AsnSerIleArgHisAsnLeuSerIleHisSerArgPheMetArgIleGlnAsnIugly 80
 QY 907 AACTCGATCCGTCACATCTGTCATCTGTCATCTGTCATCTGTCATCTGTCATCTG 966

81 AlaGlyLysserSerTyrTrpValIleAsnProAspDalaIysProGlyMetAsnProArg 100
 Db 967 GCGGAAAGAGCTCGTCGTCGTTATTAAATCCAGATGCAARGCCACGGAAATCCAGG 1026

101 ArgThrArg 103
 Db 1027 CGTACACGT 1035

RESULT 2
 US-09-645-629-22
 ; Sequence 22, Application US/09645629
 ; Patent No. 6472515
 ; GENERAL INFORMATION:
 ; APPLICANT: Clement-Johansson, Isabel
 ; APPLICANT: Dahlman-Wright, Karin
 ; APPLICANT: Lake, Staffan
 ; APPLICANT: Wasserman, Wyeth
 ; TITLE OF INVENTION: NOVEL RESPONSE ELEMENT
 ; FILE REFERENCE: 114425-032001
 ; CURRENT APPLICATION NUMBER: US/09/645-629
 ; PRIOR APPLICATION NUMBER: SE 9904269-9
 ; PRIOR FILING DATE: 1999-11-25
 ; PRIOR APPLICATION NUMBER: US 60/151,867
 ; PRIOR FILING DATE: 1999-08-31
 ; PRIOR APPLICATION NUMBER: SE 9903009-0
 ; PRIOR FILING DATE: 1999-08-26
 ; NUMBER OF SEQ ID NOS: 34
 ; SOFTWARE: FastSEQ for Windows Version 4.0
 ; SEQ ID NO: 22
 ; LENGTH: 1073
 ; TYPE: DNA
 ; ORGANISM: Artificial Sequence
 ; FEATURE:
 ; OTHER INFORMATION: synthetically generated GST-AfxDBD construct
 ; US-09-645-629-22

Alignment Scores:
 Pred. No.: 3.4e-44 Length: 1073
 Score: 410.50 Matches: 73
 Percent Similarity: 86.27% Conservative: 15
 Best Local Similarity: 71.57% Mismatches: 13
 Query Match: 73.96% Indels: 1
 DB: 4 Gaps: 1

US-09-844-353A-54 (1-103) × US-09-645-629-22 (1-1073)

1 LysLysThrThrArgArgAsnAlaTyrGlyAsnMetSerTyrAlaGluLeuIleThr 20
 QY 715 CGGAAAGGGGGCTCCGCCGAAATGCCGAAATCAGTCATCGAG 774

21 ThrlAlaIleMetAlaSerProGluLySArgLeuIleLeuAlaGlnValItyGluTrpMet 40
 Db 775 CAGGCCATGAAAGGCCCGGAGAAGCGACTCTCCACATCTACGAGTGATG 834

41 ValGlnAsnValProtYrPheArgAspLysGlyAspSerAsnSerAlaGlyTrpLys 60
 Qy 41 ValGlnAsnValProtYrPheArgAspLysGlyAspSerAsnSerAlaGlyTrpLys 60

Db 835 GTCCGTACTGTACCTACTCTAACGAAACGGTGACAGCAACGCTCAGCAGGTGAAAG 894
 Qy 61 AsnSerIleArgHisAsnLeuSerIleHisSerArgPheMetArgIleGlnAsnIugly 80
 Db 895 AACGTGATCCGCCAACCTGTCCTGCACAGTTCATCAAGTCAGGGCC 954

81 AlaGlyLysserSerTyrTrpValIleAsnProAsp -- AlanlysProGlyMetAsnPro 99
 Db 955 ACCGCGAAANGCTTGGGSGATGTTGACCTGGGAGGGCAAGGCACAGGCACAAAGCCCC 1014

RESULT 3
 US-09-620-312D-314
 ; Sequence 314, Application US/09620312D
 ; Patent No. 6569662
 ; GENERAL INFORMATION:
 ; APPLICANT: Tang, Y. Tom
 ; APPLICANT: Liu, Chenghua
 ; APPLICANT: Asundi, Vinod
 ; APPLICANT: Zhang, Jie
 ; APPLICANT: Ren, Feiyan
 ; APPLICANT: Chen, Rui-hong
 ; APPLICANT: Zhao, Qing A.
 ; APPLICANT: Wehrman, Tom
 ; APPLICANT: Xue, Aidong J.
 ; APPLICANT: Yang, Yonghong
 ; APPLICANT: Wang, Jian-Rui
 ; APPLICANT: Zhou, Ping
 ; APPLICANT: Ma, Yunqing
 ; APPLICANT: Wang, Dunru
 ; APPLICANT: Wang, Zhiwei
 ; APPLICANT: John Tillingshaft
 ; APPLICANT: Drmanac, Radivoje T.
 ; TITLE OF INVENTION: No. 6569662el Nucleic Acids and
 ; POLypeptides
 ; FILE REFERENCE: 784CIPB
 ; CURRENT APPLICATION NUMBER: US/09/620-312D
 ; CURRENT FILING DATE: 2000-07-19
 ; PRIOR APPLICATION NUMBER: 09/1552.317
 ; PRIOR FILING DATE: 2000-04-25
 ; PRIOR APPLICATION NUMBER: 09/1488,725
 ; PRIOR FILING DATE: 2000-01-21
 ; NUMBER OF SEQ ID NOS: 1105
 ; SOFTWARE: pt_PLI_genes Version 1.0
 ; SEQ ID NO: 314
 ; LENGTH: 3394
 ; TYPE: DNA
 ; ORGANISM: Homo sapiens
 ; FEATURE:
 ; NAME/KEY: CDS
 ; LOCATION: (4117)..(1934)
 ; DB: US-09-620-312D-314

Alignment Scores:
 Pred. No.: 1.71e-43 Length: 3394
 Score: 410.50 Matches: 73
 Percent Similarity: 86.27% Conservative: 15
 Best Local Similarity: 71.57% Mismatches: 13
 Query Match: 73.96% Indels: 1
 DB: 4

US-09-844-353A-54 (1-103) × US-09-620-312D-314 (1-3394)

QY 1 LysLysThrThrArgArgAsnAlaTyrGlyAsnMetSerTyrAlaGluLeuIleThr 20
 Db 690 CGGAGGGGGCTCCGCCGAAATGCCGAAATCAGTCATCGAG 749

21 ThrlAlaIleMetAlaSerProGluLySArgLeuIleLeuAlaGlnValItyGluTrpMet 40
 Qy 21 ThrlAlaIleMetAlaSerProGluLySArgLeuIleLeuAlaGlnValItyGluTrpMet 40

Db 750 CAGGCCATGAAAGGCCCGGAGAAGCGACTCTCCACATCTACGAGTGATG 809

QY 41 ValGlnAsnValProTyroPheArgAspLysGlyAspSerAsnSerAlaGlyTriPhe 60
 Db 810 GTCGGTACTGACCTACTCAAGAACAGTGACAGTCAGGATGGAA 869

QY 61 AsnSerIleArgHisAsnLeuSerLeuHisSerArgPheMetArgIleGlnAsnGluGly 80
 Db 870 AACTCGATGCCAACACCTGTCCCTGCACAGCAAGTTCATCAAGGTTACAACGAGGCC 929

QY 81 AlaGlyLysSerSerItpTrpValleAsnProAsp---AlaLysProGlyMetAsnPro 99
 Db 930 ACCGSCAAGAACGCTCTGGGATGCTGAACCTGAGGGCAGAGGCCAAAGCCCC 989

QY 100 ArgArg 101
 Db 990 CGCCGC 995

RESULT 4
 US-08-857-076-44
 ; Sequence 44, Application US/08857076C

; GENERAL INFORMATION:
 ; Patent No. 6225120

; APPLICANT: Ruvkin, Gary
 ; APPLICANT: Kimura, Routarou
 ; APPLICANT: Patterson, Garth
 ; APPLICANT: Paridis, Suzanne
 ; APPLICANT: Tissenbaum, Heidi
 ; APPLICANT: Morris, Jason
 ; APPLICANT: Kowek, Allison
 ; TITLE OF INVENTION: IMPAIRED GLUCOSE TOLERANCE CONDITIONS
 ; TITLE OF INVENTION: IMPAIRED GLUCOSE TOLERANCE CONDITIONS
 ; TITLE REFERENCE: 00736/351/001
 ; CURRENT APPLICATION NUMBER: US/08/857,076C
 ; CURRENT FILING DATE: 1997-05-15
 ; NUMBER OF SEQ ID NOS: 114
 ; SOFTWARE: FastSEQ for Windows Version 4.0
 ; SEQ ID NO 44
 ; LENGTH: 2704
 ; TYPE: DNA
 ; ORGANISM: Caenorhabditis elegans
 US-08-857-076-44

Alignment Scores:
 Prod. No.: 1.87e-39
 Score: 2704
 Percent Similarity: 75.93%
 Best Local Similarity: 62.04%
 Query Match: 68.20%
 DB: 3

US-09-844-353A-54 (1-103) x US-08-857-076-44 (1-2704)

QY 1 LysLysThrThr-----ThrArgArgAsnAlaItpGlyAsnMetSerTerY 15
 Db 398 AAGAACGCCAACCGATCATATTGGCACAGAACGATCATGGGGATCCATCTT 457

QY 16 AlaGlutLeuIleThrThrAlaIleMetAlaSerProGluLysArgLeuThrLeuAlaGln 35
 Db 458 TCGGATACTATTGCCAACGATGGAAATCGGCCAGAGGGCTTAACTCATGAG 517

QY 36 ValTyrglutItpMetValGlnAsnValProTyroPheArgAspLysGlyAspSerAsnSer 55
 Db 518 ATTATCATGGTCTCTGATATATTCCTACTTGGAAAGATCTGGTCCGAGGAG 577

QY 56 SerAlaGlyTrpLysSerAsnSerIleArgHisAsnLeuSerLeuHisSerArgPheMetArg 75
 Db 578 GCCGCCGGATGGAAATCGATCCGTCATTCCTGTTCACTCGCA 637

QY 76 IleGlnAsnGluGlyLysSerSerTrpIvalIleAsnProAspAlaLysPro 95
 Db 638 ATTAGATGAAAGAGGCCGGAAAGAGCTCGTGTGGTTAAATCCAGATCAAAGCCA 697

QY 96 GlyMetAsnProArgArgGlyArg 103
 Db 698 GGAAATGAAATCACGGCGTACAGT 721

RESULT 5
 US-07-882-292-1
 ; Sequence 1, Application US/078822292
 ; Patent No. 5324618

; GENERAL INFORMATION:
 ; APPLICANT: Tao, Wufan
 ; APPLICANT: Lai, Eseng
 ; TITLE OF INVENTION: BRAIN TRANSCRIPTION FACTOR, NUCLEIC ACIDS
 ; NUMBER OF SEQUENCES: 2

; CORRESPONDENCE ADDRESS:
 ; ADDRESSEE: John P. White
 ; STREET: c/o Cooper and Dunham, 30 Rockefeller
 ; CITY: Plaza
 ; STATE: New York
 ; COUNTRY: USA
 ; ZIP: 10112

; COMPUTER READABLE FORM:
 ; MEDIUM TYPE: Floppy disk
 ; COMPUTER: IBM PC compatible
 ; OPERATING SYSTEM: PC-DOS/MS-DOS
 ; SOFTWARE: PatentIn Release #1.24
 ; CURRENT APPLICATION DATA:
 ; APPLICATION NUMBER: US/07/882,292
 ; FILING DATE: 1992/05/13
 ; CLASSIFICATION: 514
 ; ATTORNEY/AGENT INFORMATION:
 ; NAME: White, John P
 ; REGISTRATION NUMBER: 2B, 678
 ; REFERENCE/DOCKET NUMBER: 414472
 ; TELECOMMUNICATION INFORMATION:
 ; TELEPHONE: 212-977-9550
 ; TELEFAX: 212-664-0525
 ; TELEX: 422523 COOP UT
 ; INFORMATION FOR SEQ ID NO: 1:
 ; SEQUENCE CHARACTERISTICS:
 ; LENGTH: 2830 base pairs
 ; TYPE: NUCLEIC ACID
 ; STRANDEDNESS: single
 ; TOPOLOGY: linear
 ; MOLECULE TYPE: cDNA
 ; HYPOTHETICAL: N
 ; ANTI-SENSE: N
 ; FEATURE:
 ; NAME/KEY: CDS
 ; LOCATION: 443..1882
 ; OTHER INFORMATION:
 ; NAME/KEY: misc_feature
 ; LOCATION: 926..1255
 ; OTHER INFORMATION: /note= "nucleotide sequence encoding DNA
 ; binding domain homology"
 ; FEATURE:
 ; NAME/KEY: misc_signal
 ; LOCATION: 1883..1885
 ; OTHER INFORMATION: /note= "translation termination codon"
 ; US-07-882-292-1

Alignment Scores:
 Pred. No.: 3.11e-14
 Score: 185.50
 Percent Similarity: 68.29%
 Best Local Similarity: 42.68%
 Query Match: 33.42%
 DB: 1
 Length: 2830
 Matches: 35
 Conservative: 21
 Mismatches: 19
 Indels: 7
 Gaps: 2

US-09-844-353A-54 (1-103) x US-07-882-292-1 (1-2830)

QY 14 SerTyrAlaGluLeuIleThrThrAlaIleMetAlaSerProGluLysArgLeuThrLeu 33
 :|||:|||:|||:|||:|||:|||:|||:|||:|||:|||:|||:|||:|||:|||:|||:
 Db 968 ACCTATAACGGCTCATAGATGCCATAGGGAGAGTCGGAGAARGCCCTAACGTC 1027
 ;|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:
 QY 34 AlaGlnValtryGluTrpMetValGlnAsnValProTyPheArgAspLysGlyAspSer 53
 ::|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:
 Db 1028 AACGGCTATCAGAGTTCATGAGAACATTCCTACTACCGCGAG----- 1075
 ::|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:
 QY 54 AsnSerSerAlaGlyTrpLysAsnSerTyrLargHISBasmLeuSerLeuHisSerArgPhe 73
 ::|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:
 Db 1076 ---AACAGGAGGGCTGGCAACTCCATGGCCACAACCTGTCCCTCARCAAGTCGTC 1132
 ::|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:
 QY 74 MetArgIleGlnAsn----GluGlyAlaLacLysserSerTyrPheValleAsnPro 91
 ::|:|:|:|:|:|:|:|:|:|:|:|:
 Db 1133 GTGAGGTGTCGGCCGACTACGACGACCCGGCAACTATGAGGCTGACCCG 1192
 ::|:|:|:|:|:|:|:
 QY 92 AspAla 93
 ::|:
 Db 1193 TCGAGC 1198
 ::|:
 RESULT 6
 US-08-331-644-1
 Sequence 1, Application US/08331644
 ;|:|:|:|:|:|:
 ;|:|:|:|:|:|:
 ;|:|:|:|:|:
 ;|:|:|:|:
 ;|:|:|:
 ;|:|:
 ;|:
 ;:
 GENERAL INFORMATION:
 ;|:
 APPLICANT: Tao, Wufan
 ;|:
 APPLICANT: Lai, Eseng
 ;|:
 TITLE OF INVENTION: BRAIN TRANSCRIPTION FACTOR, NUCLEIC
 ;|:
 NUMBER OF INVENTION: ACIDS ENCODING SAME AND USES THEREOF
 ;|:
 NUMBER OF SEQUENCES: 4
 ;|:
 CORRESPONDENCE ADDRESS:
 ;|:
 ADDRESSEE: Cooper & Dunham
 ;|:
 STREET: 1185 Avenue of the Americas
 ;|:
 CITY: New York
 ;|:
 STATE: New York
 ;|:
 COUNTRY: USA
 ;|:
 ZIP: 10036
 ;|:
 COMPUTER READABLE FORM:
 ;|:
 MEDIUM TYPE: Floppy disk
 ;|:
 COMPUTER: IBM PC compatible
 ;|:
 OPERATING SYSTEM: PC-DOS/MS-DOS
 ;|:
 SOFTWARE: Patent In Release #1.0, Version #1.25
 ;|:
 CURRENT APPLICATION DATA:
 ;|:
 APPLICATION NUMBER: US/08/331,644
 ;|:
 FILING DATE:
 ;|:
 CLASSIFICATION: 435
 ;|:
 PRIORITY DATA:
 ;|:
 APPLICATION NUMBER: US 07/882,292
 ;|:
 TELECOMMUNICATION INFORMATION:
 ;|:
 TELEPHONE: 212-278-0400
 ;|:
 TELEFAX: 212-291-0225
 ;|:
 INFORMATION FOR SEQ ID NO: 1:
 ;|:
 SEQUENCE CHARACTERISTICS:
 ;|:
 LENGTH: 2830 base pairs
 ;|:
 TYPE: nucleic acid
 ;|:
 STRANDEDNESS: single
 ;|:
 TOPOLOGY: linear
 ;|:
 MOLECULE TYPE: cDNA
 ;|:
 HYPOTHETICAL: N
 ;|:
 ANTI-SENSE: N
 ;|:
 FEATURE:
 ;|:
 NAME/KEY: CDS
 ;|:
 LOCATION: 443..1882
 ;|:
 OTHER INFORMATION:
 ;|:
 NAME/KEY: misc_feature
 ;|:
 LOCATION: 926..1255
 ;|:
 OTHER INFORMATION: /note= "nucleotide sequence"

;|:
 OTHER INFORMATION: encoding DNA binding domain
 ;|:
 OTHER INFORMATION: homology"
 ;|:
 FEATURE:
 ;|:
 NAME/KEY: misc_signal
 ;|:
 LOCATION: 1883..1885
 ;|:
 OTHER INFORMATION: /note= "translation termination codon"
 ;|:
 OTHER INFORMATION: codon"
 ;|:
 US-08-331-644-1
 ;|:
 Alignment Scores:
 ;|:
 Pred. No.: 3.16e-14
 ;|:
 Score: 185.50
 ;|:
 Percent Similarity: 68.29%
 ;|:
 Best Local Similarity: 42.68%
 ;|:
 Query Match: 33.42%
 ;|:
 DB: 2
 ;|:
 Gaps: 2
 ;|:
 US-09-844-353A-54 (1-103) x US-08-331-644-1 (1-2830)
 ;|:
 Qy 14 SerTyrAlaGluLeuIleThrThrAlaIleMetAlaSerProGluLysArgLeuThrLeu 33
 ;|:
 ::|:|:|:|:|:|:|:|:|:|:|:|:|:|:
 Db 968 ACCTATAACGGCTCATAGATGCCATAGGGAGAGTCGGAGAARGCCCTAACGTC 1027
 ;|:
 ::|:|:|:|:|:|:|:|:|:|:|:
 Qy 34 AlaGlnValtryGluTrpMetValGlnAsnValProTyPheArgAspLysGlyAspSer 53
 ;|:
 ::|:|:|:|:|:|:|:|:|:|:
 Db 1028 AACGGCTATCAGAGTTCATGAGAACATTCCTACTACCGCGAG----- 1075
 ;|:
 ::|:|:|:|:|:|:
 Qy 54 AsnSerSerAlaGlyTrpLysAsnSerTyrLargHISBasmLeuSerLeuHisSerArgPhe 73
 ;|:
 ::|:|:|:|:|:|:
 Db 1076 ---AACAGGAGGGCTGGCAACTCCATGGCCACAACCTGTCCCTCARCAAGTCGTC 1132
 ;|:
 ::|:|:|:|:|:
 Qy 74 MetArgIleGlnAsn----GluGlyAlaLacLysserSerTyrPheValleAsnPro 91
 ;|:
 ::|:|:|:|:
 Db 1133 GTGAGGTGTCGGCCGACTACGACGACCCGGCAACTATGAGGCTGACCCG 1192
 ;|:
 ::|:|:|:
 Qy 92 AspAla 93
 ;|:
 Db 1193 TCGAGC 1198
 ;|:
 RESULT 7
 PCT-US93-04102-1
 Sequence 1, Application PC/TUSS9304102
 ;|:
 GENERAL INFORMATION:
 ;|:
 APPLICANT: Tao, Wufan
 ;|:
 APPLICANT: Lai, Eseng
 ;|:
 TITLE OF INVENTION: BRAIN TRANSCRIPTION FACTOR, NUCLEIC
 ;|:
 ACIDS ENCODING SAME AND USES THEREOF
 ;|:
 NUMBER OF SEQUENCES: 4
 ;|:
 CORRESPONDENCE ADDRESS:
 ;|:
 ADDRESSEE: John P. White
 ;|:
 STREET: c/o Cooper and Dunham, 30 Rockefeller Plaza
 ;|:
 CITY: New York
 ;|:
 STATE: New York
 ;|:
 COUNTRY: USA
 ;|:
 ZIP: 10112
 ;|:
 COMPUTER READABLE FORM:
 ;|:
 MEDIUM TYPE: Floppy disk
 ;|:
 COMPUTER: IBM PC compatible
 ;|:
 OPERATING SYSTEM: PC-DOS/MS-DOS
 ;|:
 SOFTWARE: Patent In Release #1.0, Version #1.25
 ;|:
 CURRENT APPLICATION DATA:
 ;|:
 APPLICATION NUMBER: PCT/US93/04102
 ;|:
 FILING DATE: 13-MAY-1992
 ;|:
 ATTORNEY/AGENT INFORMATION:
 ;|:
 NAME: White, John P.
 ;|:
 REGISTRATION NUMBER: 28,678
 ;|:
 TELECOMMUNICATION INFORMATION:
 ;|:
 TELEPHONE: 212-278-0400
 ;|:
 FAX: 212-291-0225
 ;|:
 INFORMATION FOR SEQ ID NO: 1:
 ;|:
 SEQUENCE CHARACTERISTICS:
 ;|:
 LENGTH: 2830 base pairs
 ;|:
 TYPE: nucleic acid
 ;|:
 STRANDEDNESS: single
 ;|:
 TOPOLOGY: linear
 ;|:
 MOLECULE TYPE: cDNA
 ;|:
 HYPOTHETICAL: N
 ;|:
 ANTI-SENSE: N
 ;|:
 FEATURE:
 ;|:
 NAME/KEY: CDS
 ;|:
 LOCATION: 443..1882
 ;|:
 OTHER INFORMATION:
 ;|:
 NAME/KEY: misc_feature
 ;|:
 LOCATION: 926..1255
 ;|:
 OTHER INFORMATION: /note= "nucleotide sequence"

TELEFAX: 212-664-0525
 TELEX: 422523 COOP UI
 INFORMATION FOR SEQ ID NO: 1:
 SEQUENCE CHARACTERISTICS:
 LENGTH: 2830 base pairs
 TYPE: NUCLEIC ACID
 STRANDEDNESS: single
 TOPOLOGY: linear
 MOLECULE TYPE: cDNA
 HYPOTHETICAL: N
 ANTI-SENSE: N
 FEATURE:
 NAME/KEY: CDS
 LOCATION: 443..1882
 OTHER INFORMATION:
 FEATURE:
 NAME/KEY: misc_feature
 LOCATION: 926..1255
 OTHER INFORMATION: /note= "nucleotide sequence encoding DNA
 binding domain homology"
 OTHER INFORMATION:
 FEATURE:
 NAME/KEY: misc_signal
 LOCATION: 1883..1885
 OTHER INFORMATION: /note= "translation termination codon"
 PCT-US93-04102-1

Alignment Scores:

Pred. No.:	Score:	Percent Similarity:	Best Local Similarity:	Query Match:	DB:
	3.16e-14	185.50	68.29%	33.42%	
					US-09-844-353A-54 (1-103) x US-09-331-644-3 (1-1860)
					US-09-844-353A-54 (1-103) x US-09-331-644-3 (1-1860)
					US-09-844-353A-54 (1-103) x US-09-331-644-3 (1-1860)

Prod. No.:
 Score: 35
 Percent Similarity: 21
 Best Local Similarity: 19
 Query Match: 7
 DB: 5

Length: 2830
 Matches: 35
 Conservative: 21
 Mismatches: 19
 Indels: 7
 Gaps: 2

Length: 1860
 Score: 184.50
 Percent Similarity: 67.07%
 Best Local Similarity: 43.90%
 Query Match: 33.24%
 DB: 2

Length: 36
 Matches: 36
 Conservative: 19
 Mismatches: 20
 Indels: 7
 Gaps: 2

Length: 875
 Score: 2.37e-14
 Percent Similarity: 6.78%
 Best Local Similarity: 4.35%
 Query Match: 2.35%
 DB: 3

Length: 875
 Score: 184.50
 Percent Similarity: 67.07%
 Best Local Similarity: 43.90%
 Query Match: 33.24%
 DB: 2

Length: 73
 Score: 14
 Percent Similarity: 5.71%
 Best Local Similarity: 4.35%
 Query Match: 2.35%
 DB: 3

Length: 73
 Score: 765
 Percent Similarity: 67.07%
 Best Local Similarity: 43.90%
 Query Match: 33.24%
 DB: 2

Length: 53
 Score: 34
 Percent Similarity: 6.78%
 Best Local Similarity: 4.35%
 Query Match: 2.35%
 DB: 3

Length: 53
 Score: 825
 Percent Similarity: 6.78%
 Best Local Similarity: 4.35%
 Query Match: 2.35%
 DB: 3

Length: 73
 Score: 54
 Percent Similarity: 6.78%
 Best Local Similarity: 4.35%
 Query Match: 2.35%
 DB: 3

Length: 73
 Score: 876
 Percent Similarity: 6.78%
 Best Local Similarity: 4.35%
 Query Match: 2.35%
 DB: 3

Length: 91
 Score: 74
 Percent Similarity: 6.78%
 Best Local Similarity: 4.35%
 Query Match: 2.35%
 DB: 3

Length: 91
 Score: 930
 Percent Similarity: 6.78%
 Best Local Similarity: 4.35%
 Query Match: 2.35%
 DB: 3

Length: 93
 Score: 92
 Percent Similarity: 6.78%
 Best Local Similarity: 4.35%
 Query Match: 2.35%
 DB: 3

Length: 95
 Score: 990
 Percent Similarity: 6.78%
 Best Local Similarity: 4.35%
 Query Match: 2.35%
 DB: 3

RESULT 9
 PCT-US93-04102-3
 Sequence 3, Application PC/TUS9304102
 General Information:
 Applicant: Tao, Wufan
 Applicant: Lai, Eseng
 Title of Invention: BRAIN TRANSCRIPTION FACTOR, NUCLEIC
 Number of Sequences: 4
 Correspondence Address:
 Addressee: Cooper & Dunham
 Street: 1185 Avenue of the Americas
 City: New York
 State: New York
 Country: USA
 ZIP: 10036

RESULT 9
 PCT-US93-04102-3
 Sequence 3, Application PC/TUS9304102
 General Information:
 Applicant: Tao, Wufan
 Applicant: Lai, Eseng
 Title of Invention: BRAIN TRANSCRIPTION FACTOR, NUCLEIC
 Number of Sequences: 4
 Correspondence Address:
 Addressee: John P. White
 Street: c/o Cooper and Dunham, 30 Rockefeller Plaza
 City: New York
 State: New York
 Country: USA

ZIP: 10112 NUMBER OF SEQUENCES: 22
 COMPUTER READABLE FORM:
 i MEDIUM TYPE: Floppy disk
 COMPUTER: IBM PC compatible
 OPERATING SYSTEM: PC-DOS/MS-DOS
 SOFTWARE: PatentIn Release #1.0, Version #1.125
 CURRENT APPLICATION DATA:
 APPLICATION NUMBER: PCT/US93/04102
 FILING DATE: 19930430
 CLASSIFICATION:
 PRIORITY APPLICATION DATA:
 APPLICATION NUMBER: US 07/882,292
 FILING DATE: 13-MAY-1992
 ATTORNEY/AGENT INFORMATION:
 NAME: White, John P.
 REGISTRATION NUMBER: 28,678
 REFERENCE/DOCKET NUMBER: 41472A-PCT
 TELECOMMUNICATION INFORMATION:
 TELEPHONE: 212-977-9550
 TELEFAX: 212-664-0525
 INFORMATION FOR SEQ ID NO: 3:
 LENGTH: 1860 base pairs
 TYPE: NUCLEIC ACID
 STRANDEDNESS: single
 TOPOLOGY: Linear
 MOLECULE TYPE: cDNA
 PCT-US93-04102-3

Alignment scores:
 Pred. No.: 2.37e-14 Length: 1860
 Score: 184.50 Matches: 36
 Percent Similarity: 67.07% Conservative: 19
 Best Local Similarity: 43.90% Mismatches: 20
 Query Match: 33.24% Irrels: 7
 DB: 5 Gaps: 2

US-09-844-353A-54 (1-103) × PCT-US93-04102-3 (1-1860)

Qy 14 SerTyrAlaGluLeuIleThrThrAlaIleMetAlaSerProGluLysArgLeuThrLeu 33
 Db 765 TGTGTAATCGCGCTCATCACCATGGCCATCCTGCAGGCCAAAGAAGGCCGCTGACCTC 824

Qy 34 AlaGlnValtyrGluTrpMetValGlnAsnValProtYrPheArgAspLysGlyAspSer 53
 Db 825 AGGGAGCTGCGAACGTCATCAGCGGCTTCCTACTACCGGAG---- 875

Qy 54 AsnSerSerAlaGlyTrpLysAsnSerIleArgHisAsnLeuSerIleHisSerArgPhe 73
 Db 876 -----TTCCCCCTGCGAGACAGATCCACACCTTGCTCAACCACTGTGTC 929

Qy 74 MetArgIleGlnAsnGlu-----GlyAlaGlyLysSerSerTrpTrpValIleAsnPro 91
 Db 930 GTCGAAGATCCCGCGCAACCGGCCAAAGGCCAAACTACGCTGACGCCG 989

Qy 92 AspAla 93
 Db 990 GAGTCC 995

RESULT 10
 US-09-083-351-3
 Sequence 3, Application US/09083351
 GENERAL INFORMATION:
 Patent No. 6087107
 APPLICANT: Sheffield, Val C.
 APPLICANT: Alward, Wallace L.M.
 APPLICANT: Stone, Edwin M.
 APPLICANT: Nishimura, Darryl
 APPLICANT: Patil, Shiva
 TITLE OF INVENTION: THERAPEUTICS AND DIAGNOSTICS FOR CONGENITAL HEART DISEASE BASED ON A NOVEL HUMAN TRANSCRIPTION FACTOR

RESULT 11
 US-09-083-352-3
 Sequence 3, Application US/09083352
 GENERAL INFORMATION:
 Patent No. 620450
 APPLICANT: Sheffield, Val C.
 APPLICANT: Alward, Wallace L.M.
 APPLICANT: Stone, Edwin M.
 APPLICANT: Nishimura, Darryl
 APPLICANT: Patil, Shiva

TITLE OF INVENTION: GLAUCOMA THERAPEUTICS AND DIAGNOSTICS
 NUMBER OF INVENTION: BASED ON A NOVEL HUMAN TRANSCRIPTION FACTOR
 NUMBER OF SEQUENCES: 22
 CORRESPONDENCE ADDRESS:
 ADDRESSEE: FOLEY, HOAG & ELIOT LLP
 STREET: One Post Office Square
 CITY: Boston
 STATE: MA
 COUNTRY: USA
 ZIP: 02109-2170
 COMPUTER READABLE FORM:
 MEDIUM TYPE: Floppy disk
 COMPUTER: IBM PC compatible
 OPERATING SYSTEM: PC-DOS/MS-DOS
 SOFTWARE: PatentIn Release #1.0, Version #1.30
 CURRENT APPLICATION DATA:
 APPLICATION NUMBER: US/09/083 , 352
 FILING DATE: 22-MAY-1998
 CLASSIFICATION: 514
 ATTORNEY/AGENT INFORMATION:
 NAME: Arnold, Beth E.
 REGISTRATION NUMBER: 35,430
 REFERENCE/DOCKET NUMBER: UIA-029.01
 TELECOMMUNICATION INFORMATION:
 TELEPHONE: 617-832-1000
 TELEFAX: 617-832-7000
 INFORMATION FOR SEQ ID NO: 3:
 SEQUENCE CHARACTERISTICS:
 LENGTH: 1659 base pairs
 TYPE: nucleic acid
 STRANDEDNESS: single
 TOPOLOGY: linear
 MOLECULE TYPE: cDNA
 US-09-083-352-3

Alignment Scores:
 Pred No.: 3.68e-14 Length: 1659
 Score: 182.50 Matches: 33
 Percent Similarity: 68.29% Conservative: 23
 Best Local Similarity: 40.24% Mismatches: 19
 Query Match: 32.88% Indels: 7
 DB: 3 Gaps: 2

US-09-844-353A-54 (1-103) x US-09-083-352-3 (1-1659)

Qy 14 SerTyrAlaGluLeuIleThrAlaLeuSerProGluLysArgLeuThrLeu 33
 Db 244 AGCTACATCGCCTCATCCATGCCATTGCCATGCCAACAGATCACCTG 303
 Qy 34 AlaGlnValTyrGluTrpMetValGlnAsnValProTyrPheArgAspSer 53
 Db 304 AACGGCATCTACCAAGTTATCATGGACGCCCTCCCTCTACGGGAC----- 351
 Qy 54 AsnSerSerAlaGlyTrpLysAsnSerIleArgHisAsnLeuSerArgPhe 73
 Db 352 --AACANGCAGGTGGAGAACGATCCAGCTACTGGACCTGGACCCG 408
 Qy 74 MetArgIle-----GlnAsnGluGlyAlaGlyLysSerSerTrpTrpValLeAsnPro 91
 Db 409 GTCAAGGTGCCGCGCAAGCAGAAAGCCGGAAAGGCCAGCTACTGGACCTGGACCCG 468
 Qy 92 AspAla 93
 Db 469 GACTCC 474

RESULT 12
 US 09-083-351-1
 GENERAL INFORMATION
 Sequence 1, Application US/09083351
 Patent No. 6087107
 APPLICANT: Sheffield, Val C.
 APPLICANT: Alward, Wallace L.M.
 APPLICANT: Stone, Edwin M.

Qy 74 MetArgIle-----GlnAsnGluGlyAlaGlyLysSerSerTrpTrpValLeAsnPro 91
 Db 883 GTCAAGGTGCCGCGCAAGCAGAAAGCCGGAAAGGCCAGCTACTGGACCTGGACCCG 942
 Qy 92 AspAla 93
 Db 943 GACTCC 948

RESULT 13
 US 09-083-352-1

Sequence 1, Application US/09083352
 GENERAL INFORMATION:
 APPLICANT: Sheffield, Val C.
 APPLICANT: Alward, Wallace L.M.
 APPLICANT: Stone, Edwin M.
 APPLICANT: Nishimura, Darryl
 APPLICANT: Patel, Shiva
 TITLE OF INVENTION: GLAUCOMA THERAPEUTICS AND DIAGNOSTICS
 TITLE OF INVENTION: BASED ON A NOVEL HUMAN TRANSCRIPTION FACTOR
 NUMBER OF SEQUENCES: 22
 CORRESPONDENCE ADDRESS:
 ADDRESSEES: FOLEY, HOAG & ELIOT LLP
 STREET: One Post Office Square
 CITY: Boston
 STATE: MA
 COUNTRY: USA
 ZIP: 02109-2170
 COMPUTER READABLE FORM:
 MEDIUM TYPE: Floppy disk
 COMPUTER: IBM PC Compatible
 OPERATING SYSTEM: PC-DOS/MS-DOS
 SOFTWARE: PatentIn Release #1.0, Version #1.30
 CURRENT APPLICATION DATA:
 APPLICATION NUMBER: US/09/083,352
 FILING DATE: 22-MAY-1998
 CLASSIFICATION: 514
 ATTORNEY/AGENT INFORMATION:
 NAME: Arnold, Beth E.
 REFERENCE NUMBER: 35,430
 REFERENCE/DOCKET NUMBER: UIA-029-01
 TELECOMMUNICATION INFORMATION:
 TELEPHONE: 617-832-1000
 TELEFAX: 617-832-7000
 SEQUENCE CHARACTERISTICS:
 LENGTH: 946 base pairs
 TYPE: nucleic acid
 STRANDEDNESS: single
 TOPOLOGY: linear
 MOLECULE TYPE: cDNA
 FEATURE:
 NAME/KEY: CDS
 LOCATION: 475..2133
 US-09-083-352-1

Db 943 GACTCC 948
 RESULT 14
 US-08-742-753-3
 Sequence 3, Application US/08742753
 Patent No. 5861278
 GENERAL INFORMATION:
 APPLICANT: WONG, Gordon G.
 APPLICANT: YAO, KWOK-MING
 TITLE OF INVENTION: HNF3-delta Compositions
 NUMBER OF SEQUENCES: 4
 CORRESPONDENCE ADDRESS:
 ADDRESSEES: Genetics Institute, Inc.
 STREET: 87 Cambridge Park Drive
 CITY: Cambridge
 STATE: Massachusetts
 COUNTRY: USA
 ZIP: 02140
 COMPUTER READABLE FORM:
 MEDIUM TYPE: Floppy disk
 COMPUTER: IBM PC Compatible
 OPERATING SYSTEM: PC-DOS/MS-DOS
 SOFTWARE: PatentIn Release #1.0, Version #1.25
 CURRENT APPLICATION DATA:
 APPLICATION NUMBER: US/08/742,753
 FILING DATE:
 CLASSIFICATION: 514
 ATTORNEY/AGENT INFORMATION:
 NAME: LAZAR, Steven R.
 REGISTRATION NUMBER: 32,618
 REFERENCE/DOCKET NUMBER: 5277
 TELECOMMUNICATION INFORMATION:
 TELEPHONE: (617) 498-9260
 TELEFAX: (617) 876-5851
 INFORMATION FOR SEQ ID NO: 3:
 SEQUENCE CHARACTERISTICS:
 LENGTH: 3342 base pairs
 TYPE: nucleic acid
 STRANDEDNESS: single
 TOPOLOGY: linear
 MOLECULE TYPE: cDNA
 FEATURE:
 NAME/KEY: CDS
 LOCATION: 70..2358
 US-08-742-753-3

Alignment Scores:
 Pred. No.: 1.24e-13 Length: 3946
 Score: 182.50 Matches: 33
 Percent Similarity: 68.29% Conservative: 23
 Best Local Similarity: 40.24% Mismatches: 19
 Query Match: 32.88% Indels: 7
 Db: 3 Gaps: 2

US-09-084-353A-54 (1-103) x US-09-083-352-1 (1-3946)

Qy 14 SerTyrAlaGluLeuIleThrThrAlaIleMetAlaSerProGlyLysArgLeuThrLeu 33
 Db 718 AGCTACATGGCTCATCCTATGGCTTCAGCAACGCCGAGTCACCTG 777

Qy 34 AlaGlnValTyrGlutrpMetValGlnAsnValProTyprPheArgAspLysGlyAspSer 53
 Db 778 AACGGCATATCCAGTATCATCAGAACGCTTCACCTCGCTAACCTGGCAC 825

Qy 54 AsnSerSerAlaGlyTrpLysAsnSerTyrLeuSerIleHisSerArgPhe 73
 Db 826 ---AACAAAGCAGGCTGGAGAACGACATCCGCAACCTCTCGCTAACCTGGCTTC 882

Qy 74 MetArgIle-----GlnAsnGluGlyAlaGlyLysSerSerTrpPheValLeAspPro 91
 Db 883 GTCAAGGTGGCGCGAGCACAGGGCAAGGCCGAGCTACTGGACGCTGGACCG 942

Qy 92 AspAla 93
 Db 919 TCATCCGCACACTTCCCTGGACGACTGGACGCTGG 975

Qy 9 AlanrpGlyAsn-----MetserTyrAlaGluLeuIleThrLeuIleThr 21
 Db 751 TCCTGGAGAACTCTGTTGCTGAGGCCACCCACTCTTACATGGCATGAAATT 810

Qy 22 AlaIleMetAlaSerProGlyLysArgLeuThrLeuAlaGlnValTyrGluTrpMetVal 41
 Db 811 GCATTAACAGCATGAGGAGGCGACTTGAAAGACATACTAGCTGATGATGAG 870

Qy 42 GluAsnValProTyrPheArgAspLysGlyAspSerAsnSerSerAlaGlyTrpAsn 61
 Db 871 GACCACTTTCCCTACTTTAAAG----CACATTGCCAACGGAGCTGGAAAC 918

Qy 62 SerIleArgGlyAsnLeuSerIleHisSerArgPheMetArgIleGlnAsnGluGlyAla 81
 Db 919 TCATCCGCACACTTCCCTGGACGACTGGACGCTGG 975

Qy 82 GlyLysSerSerTrpTrpValLeuAsnProAspAla 93
 Db 976 GGAAAGGTCTCTTGACCATCACCCAGTGGC 1011
 RESULT 15
 US-08-331-644-4 Sequence 4, Application US/08331644
 ; Patent No. 5976872
 ; GENERAL INFORMATION:
 ; APPLICANT: Tao, Wufan
 ; TITLE OF INVENTION: BRAIN TRANSCRIPTION FACTOR, NUCLEARIC
 ; TITLE OF INVENTION: ACIDS ENCODING SAME AND USES THEREOF
 ; NUMBER OF SEQUENCES: 4
 ; CORRESPONDENCE ADDRESS:
 ; ADDRESSEE: Cooper & Dunham
 ; STREET: 1185 Avenue of the Americas
 ; CITY: New York
 ; STATE: New York
 ; COUNTRY: USA
 ; ZIP: 10036
 ; COMPUTER READABLE FORM:
 ; MEDIUM TYPE: FLOPPY disk
 ; COMPUTER: IBM PC compatible
 ; OPERATING SYSTEM: PC-DOS/MS-DOS
 ; SOFTWARE: Patent In Release #1.0, Version #1.25
 ; CURRENT APPLICATION DATA:
 ; APPLICATION NUMBER: US/08/331,644
 ; FILING DATE:
 ; CLASSIFICATION: 435
 ; PRIOR APPLICATION DATA:
 ; APPLICATION NUMBER: US 07/882,292
 ; FILING DATE: 13-MAY-1992
 ; ATTORNEY/AGENT INFORMATION:
 ; NAME: White, John P.
 ; REGISTRATION NUMBER: 28,678
 ; REFERENCE DOCKET NUMBER: 41472-A-PCT-US
 ; TELECOMMUNICATION INFORMATION:
 ; TELEPHONE: 212-278-0400
 ; TELEFAX: 212-331-0525
 ; INFORMATION FOR SEQ ID NO: 4:
 ; SEQUENCE CHARACTERISTICS:
 ; LENGTH: 1155 base pairs
 ; TYPE: nucleic acid
 ; STRANDEDNESS: single
 ; TOPOLOGY: linear
 ; MOLECULE TYPE: cDNA
 ; US-08-331-644-4

Alignment Scores:
 Pred. No.: 3e-14 Length: 1155
 Score: 181.50 Matches: 33
 Percent Similarity: 68.29% Conservative: 23
 Best Local Similarity: 40.24% Mismatches: 19
 Query Match: 32.70% Index: 7
 DB: 2 Gaps: 2

US-09-844-353A-54 (1-103) x US-08-331-644-4 (1-1155)

Qy 14 SerTerAlaGluIleUleThrThrAlaIleMetAlaSerProGluLysArgLeuThrLeu 33
 Db 511 AGCTATATAGCGCTCACTACATGGGAAACAGCATTGGCAGAACAGATACTCTG 570
 Qy 34 AlaGlnValItyrGluTrpMetValGlnAsnValProTyzPheArgAspLysGlyAspSer 53
 Db 571 AACGGCATCTACCGTTCACTGACCGTTCCCTTCACGGAG----- 618

Qy 54 AsnSerSerSerAlaGlyTrpLysAsnSerIleArgHisAsnLeuSerLeuHisSerArgPhe 73
 Db 619 ---AACAAAGCAGGCTGGCGAACAGCATTGGCAGAACACTGTCACTCAATGAGTGCCTC 675

Qy 74 MetArgLle-----GlnGlnGlyAlaGlyLysSerTerTrpValIleAsnPro 91

This Page Blank (Up to)

Copyright (c) 1993 - 2004 Compugen Ltd.

OM_protein - nucleic search, using frame_plus_p2n model

Run on: July 29, 2004, 11:31:55 : Search time 358 Seconds
(without alignments)

1410.680 Million cell updates/sec

Title: US-09-844-353A-54

Perfect score: 555

Sequence: 1 KKTTRRRNANGNMSTYALIT.....SSWWVVINPDAKPGMMPRRTR 103

Scoring table: BLOSUM62

Xgapop	10.0	Xgapext	0.5
Ygapop	10.0	Ygapext	0.5
Fgapop	6.0	Fgapext	7.0
Delop	6.0	Delext	7.0

Searched: 322919 seqs, 2451570024 residues

Total number of hits satisfying chosen parameters: 6445838

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Command line parameters:

```
-MODEL=frame_p2n.model -DEV=xlp
-Q|-cgn2_1/UPTO spoil p/us09844353/runat_23072004 084915 28361/app_query.fasta_1.263
-DB=Published Applications NA -QMTT=fastapt -SUFFIX=trpb -MINMATCH=0.1
-LOOPCL=0 -LOOPEXT=0 -UNITS=bits -START=1 -END=-1 -MATRIX=bloom62
-TRANS=human0.cdd -LIST=45 -DOCALLIGN=200 -THR SCORE=REPT -THR MAX=100
-TFR MINLEN=15 -MODE=LOCAL -OUTFORMAT=proto -NORM=ext -HEAPSIZE=500 -MINLEN=0
-MXLEN=200000000 -USR=us09844353 @cgn_1.1.912 @runat_23072004 084915_28361
-CGN2=6 -ICPBL=NO -NO MNRP -LARGEPOP -NEG SCORES=0 -WAIT -DSPBLOCK=100
-LONGLOG -DEV TIMEOUT=30 -WARN TIMEOUT=30 -THREADS=1 -XGAPPOP=1.0 -XGAPEXT=0.5
-XGAPOP=6 -PGAPEXT=7 -YGAPOB=10 -YGAPEXT=0.5 -DELEXT=7
```

Database :

Published Applications NA.*

- /cgn2_6/ptodata/1/pubna/pectus_pubcomb.seq;*
- /cgn2_6/ptodata/1/pubna/us07_pubcomb.seq;*
- /cgn2_6/ptodata/1/pubna/us06_pub_comb.seq;*
- /cgn2_6/ptodata/1/pubna/us05_pub_comb.seq;*
- /cgn2_6/ptodata/1/pubna/us04_pub_comb.seq;*
- /cgn2_6/ptodata/1/pubna/us03_pub_comb.seq;*
- /cgn2_6/ptodata/1/pubna/us02_pub_comb.seq;*
- /cgn2_6/ptodata/1/pubna/us01_pub_comb.seq;*
- /cgn2_6/ptodata/1/pubna/us09_pub_comb.seq;*
- /cgn2_6/ptodata/1/pubna/us60_pub_comb.seq;*
- /cgn2_6/ptodata/1/pubna/us60_pub_comb.seq;*
- /cgn2_6/ptodata/1/pubna/us10_pub_comb.seq;*
- /cgn2_6/ptodata/1/pubna/us10b_pub_comb.seq;*
- /cgn2_6/ptodata/1/pubna/us10c_pub_comb.seq;*
- /cgn2_6/ptodata/1/pubna/us10c_new_pub.seq;*
- /cgn2_6/ptodata/1/pubna/us09_new_pub.seq;*
- /cgn2_6/ptodata/1/pubna/us08_pub_comb.seq;*
- /cgn2_6/ptodata/1/pubna/us09_pub_comb.seq;*
- /cgn2_6/ptodata/1/pubna/us10b_pub_comb.seq;*
- /cgn2_6/ptodata/1/pubna/us10c_pub_comb.seq;*
- /cgn2_6/ptodata/1/pubna/us60_new_pub.seq;*
- /cgn2_6/ptodata/1/pubna/us60_pub_comb.seq;*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

RESULT 1

US-09-205-658-43

; Sequence 43, Application US/09205658

; Patent No. US2001029617A1

; GENERAL INFORMATION:

; APPLICANT: Ruykun, Gary

; ATTORNEY: TCGS, SECET

; TITLE OF INVENTION: THERAPEUTIC AND DIAGNOSTIC TOOLS FOR IMPAIRED GLUCOSE TOLERANCE CONDITIONS

; CURRENT APPLICATION NUMBER: US/09/205,658

; EARLIER APPLICATION NUMBER: 1998-12-03

; EARLIER FILING DATE: 1997-05-15

; EARLIER APPLICATION NUMBER: 08/888,534

; EARLIER FILING DATE: 1997-07-07

; EARLIER APPLICATION NUMBER: US98/10080

; EARLIER FILING DATE: 1998-05-15

; NUMBER OF SEQ ID NOS: 328

; SOFTWARE: FastSEQ for Windows Version 4.0

; SEQ ID NO 43

; LENGTH: 3499

; TYPE: DNA

SUMMARIES

Result No.	Score	Query	Match Length	DB ID	Description
1	555	100.0	3499	9	US-09-205-658-43
2	555	100.0	3499	9	US-09-844-353A-54
3	555	100.0	3499	10	US-09-963-693-43
4	440.5	79.4	3183	17	US-10-717-597-245
5	427.5	77.0	5723	15	US-10-007-926A-134
6	427.5	77.0	5723	16	US-10-341-434-52
7	427.5	77.0	5723	17	US-10-671-074-4
8	426.5	76.8	4945	17	US-10-671-074-1
9	410.5	74.0	1073	15	US-10-261-517-2
10	410.5	74.0	3394	15	US-10-037-270-314
11	410.5	74.0	3394	16	US-10-117-722-314
12	378.5	68.2	2704	9	US-09-205-658-44
13	378.5	68.2	2704	9	US-09-844-353A-44
14	378.5	68.2	2704	10	US-19-963-693-4
15	335.5	60.5	1353	15	US-10-186-839-1
16	189.5	34.1	834	9	US-09-764-864-35
17	189.5	34.1	834	9	US-09-764-864-35
18	188.5	34.0	2559	13	US-10-342-887-1132
19	188.5	34.0	2559	13	US-10-172-118-1132
20	188.5	34.0	2559	15	US-10-294-582-24
21	183.5	33.1	2299	9	US-09-925-300-589
22	182.5	32.9	714	15	US-10-029-386-20688
23	182.5	32.9	1662	13	US-10-092-887-611
24	182.5	32.9	1662	13	US-10-342-887-611
25	182.5	32.9	1662	13	US-10-172-118-611
26	182.5	32.9	2271	15	US-10-161-510-133
27	182.5	32.9	2487	13	US-10-342-887-1045
28	182.5	32.9	2487	13	US-10-172-118-1045
29	182.5	32.9	2487	15	US-10-205-823-135
30	182.5	32.9	5181	12	US-09-968-007A-802
31	182	32.8	2737	14	US-10-151-587-1
32	182	32.8	2737	17	US-10-161-609-1
33	182	32.8	3336	13	US-10-348-360-304
34	182	32.8	3417	15	US-10-084-817-336
35	182	32.8	3492	9	US-09-954-456-78
36	182	32.8	3492	9	US-09-954-456-712
37	182	32.8	3492	9	US-09-954-456-1820
38	182	32.8	3492	13	US-10-342-887-1887
39	182	32.8	3492	13	US-10-172-118-1887
40	182	32.8	3559	14	US-10-071-766-110
41	181.5	32.7	3559	15	US-10-084-717-337
42	181.5	32.7	15056	9	US-09-963-285-9
43	181.5	32.7	15056	17	US-10-303-635-4
44	181.5	32.7	2323	17	US-10-313-635-1
45	181.5	32.7	2712	9	US-09-963-285-6

; ORGANISM: *Caenorhabditis elegans*
US-09-205-658-43

Alignment Scores:
Pred. No.: 7.96e-70 Length: 3499
Score: 555.00 Matches: 103
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00%
Query Match: 100.00%
DB: 9 Gaps: 0

US-09-844-353A-54 (1-103) x US-09-205-658-43 (1-3499)

Qy 1 LysLysThrThrArgArgAsnAlaTrpGlyAsnMetSerThrAlaGluLeuIleThr 20
Db 727 AAAAGCAACGCCAACGGAACTGGAAATAATGTCATATGTCATACT 786

Qy 21 ThralalMetAlaSerProGluLysArgLeuThrLeuAlaGlnValTyrglurTrpMet 40
Db 787 ACAGCCATTATGGCTAGTCAGAAACGGTTAACCTTGCAACTTAACTGTGATGATG 846

Qy 41 ValGlnAsnValProTyrrpheArgAspLysGlyAspSerAsnSerSerAlaGlyTrpLys 60
Db 847 GTCCAGATGTTCACTACGGATAAGGGATTAAGGGATTCAGTCAGATGGAAG 906

Qy 61 AsnSerIleArgHisAsnLeuSerIleHisSerArgPheMetArgIleGlnAsnGluGly 80
Db 907 AAAAGCAACGCCAACGGAACTGGAAATAATGTCATATGTCATACT 786

Qy 907 AACTCGATCCTCACATCCTCTCATCTCTGTTCACTCGGTTCAATCCAGAG 906

Qy 907 ArgThrArg 1.03
Db 1027 CGTACACGT 1035

RESULT 3
US-09-963-693-43
; Sequence 43, Application US/09963693
; Publication No. US20030181364A1
; GENERAL INFORMATION:
; APPLICANT: Ruykun, Gary
; APPLICANT: Ogg, Scott
; TITLE OF INVENTION: THERAPEUTIC AND DIAGNOSTIC TOOLS FOR
; IMPAIRED GLUCOSE TOLERANCE CONDITIONS
; FILE REFERENCE: 00786/351004
; CURRENT APPLICATION NUMBER: US/09/963, 693
; CURRENT FILING DATE: 2001-09-25
; PRIOR APPLICATION NUMBER: US/09/205, 658
; PRIOR FILING DATE: 1998-12-03
; PRIOR APPLICATION NUMBER: 08/857, 076
; PRIOR FILING DATE: 1997-05-15
; PRIOR APPLICATION NUMBER: 08/888, 534
; PRIOR FILING DATE: 1997-07-07
; PRIOR APPLICATION NUMBER: US98/10080
; PRIOR FILING DATE: 1998-05-15
; NUMBER OF SEQ ID NOS: 328
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO: 43
; LENGTH: 3499
; TYPE: DNA
; ORGANISM: *Caenorhabditis elegans*
US-09-963-693-43

Alignment Scores:
Pred. No.: 7.96e-70 Length: 3499
Score: 555.00 Matches: 103
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00%
Query Match: 100.00%
DB: 9 Gaps: 0

US-09-844-353A-54 (1-103) x US-09-963-693-43 (1-3499)

Qy 1 LysLysThrThrArgArgAsnAlaTrpGlyAsnMetSerThrAlaGluLeuIleThr 20
Db 727 AAAAGCAACGCCAACGGAAATAATGTCATATGTCATACT 786

Qy 21 ThralalMetAlaSerProGluLysArgLeuThrLeuAlaGlnValTyrglurTrpMet 40
Db 787 ACAGCCATTATGGCTAGTCAGAAACGGTTAACCTTGCAACTTAACTGTGATGATG 846

Qy 41 ValGlnAsnValProTyrrpheArgAspLysGlyAspSerAsnSerSerAlaGlyTrpLys 60
Db 847 GTCCAGATGTTCACTACGGATAAGGGATTAAGGGATTCAGTCAGATGGAAG 906

Qy 61 AsnSerIleArgHisAsnLeuSerIleHisSerArgPheMetArgIleGlnAsnGluGly 80
Db 907 AAAAGCAACGCCAACGGAACTGGAAATAATGTCATATGTCATACT 786

Qy 907 AACTCGATCCTCACATCCTCTCATCTCTGTTCACTCGGTTCAATCCAGAG 906

Qy 907 ArgThrArg 1.03
Db 1027 CGTACACGT 1035

RESULT 3
US-09-963-693-43
; Sequence 43, Application US/09963693
; Publication No. US20030181364A1
; GENERAL INFORMATION:
; APPLICANT: Ruykun, Gary
; APPLICANT: Ogg, Scott
; TITLE OF INVENTION: THERAPEUTIC AND DIAGNOSTIC TOOLS FOR
; IMPAIRED GLUCOSE TOLERANCE CONDITIONS
; FILE REFERENCE: 00786/351004
; CURRENT APPLICATION NUMBER: US/09/963, 693
; CURRENT FILING DATE: 2001-09-25
; PRIOR APPLICATION NUMBER: US/09/205, 658
; PRIOR FILING DATE: 1998-12-03
; PRIOR APPLICATION NUMBER: 08/857, 076
; PRIOR FILING DATE: 1997-05-15
; PRIOR APPLICATION NUMBER: 08/888, 534
; PRIOR FILING DATE: 1997-07-07
; PRIOR APPLICATION NUMBER: US98/10080
; PRIOR FILING DATE: 1998-05-15
; NUMBER OF SEQ ID NOS: 328
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO: 43
; LENGTH: 3499
; TYPE: DNA
; ORGANISM: *Caenorhabditis elegans*
US-09-963-693-43

Alignment Scores:
Pred. No.: 7.96e-70 Length: 3499
Score: 555.00 Matches: 103
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00%
Query Match: 100.00%
DB: 9 Gaps: 0

US-09-844-353A-54 (1-103) x US-09-963-693-43 (1-3499)

Qy 1 LysLysThrThrArgArgAsnAlaTrpGlyAsnMetSerThrAlaGluLeuIleThr 20
Db 727 AAAAGCAACGCCAACGGAAATAATGTCATATGTCATACT 786

Qy 21 ThralalMetAlaSerProGluLysArgLeuThrLeuAlaGlnValTyrglurTrpMet 40
Db 787 ACAGCCATTATGGCTAGTCAGAAACGGTTAACCTTGCAACTTAACTGTGATGATG 846

Qy 41 ValGlnAsnValProTyrrpheArgAspLysGlyAspSerAsnSerSerAlaGlyTrpLys 60
Db 847 GTCCAGATGTTCACTACGGATAAGGGATTAAGGGATTCAGTCAGATGGAAG 906

Db 847 GTCCAGAANGTCCATACTTCAGGATAAGGGATTCCACAGTCAGCTGAG 906 Qy 100 ArgArg 101
Db 1666 CGGGGG 1671

RESULT 5
US-10-007-926A-134
; Sequence 134, Application US/10007926A
; Publication No. US2003013539A1
; GENERAL INFORMATION:
; APPLICANT: BEPUTUCCI, FRANCOIS
; APPLICANT: HOUGLADE, REMI
; APPLICANT: BIRNBAUM, DANIEL
; APPLICANT: NGUYEN, CATHERINE
; APPLICANT: VITENS, PATRICE
; APPLICANT: FERT, VINCENT
; TITLE OF INVENTION: GENE EXPRESSION PROFILING OF PRIMARY BREAST CARCINOMAS
; TITLE OF INVENTION: USING ARRAYS OF CANDIDATE GENES
; FILE REFERENCE: 1546-R-00
; CURRENT APPLICATION NUMBER: US/10/007,926A
; PRIORITY APPLICATION NUMBER: 60/254,090
; PRIOR FILING DATE: 2000-12-08
; NUMBER OF SEQ ID NOS: 468
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 134
; LENGTH: 5723
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; OTHER INFORMATION: forehead box ola (rhabdomyosarcoma) (FOXO1A)
; OTHER INFORMATION: gene.
US-10-007-926A-134

Alignment Scores:
Pred. No.: 5.96e-51 Length: 5723
Score: 427.50 Matches: 75
Percent Similarity: 91.18% Conservatives: 18
Best Local Similarity: 73.53% Mismatches: 8
Query Match: 77.03% Indels: 1
DB: 15 Gaps: 1

US-09-844-353A-54 (1-103) x US-10-007-926A-134 (1-5723)

Qy 1 LysLysThrThrArgAsnAlaTrpGlyAsnMetSerTyraGluLeuIleThr 20
Db 8336 ArgAGAGCTGCGCCGCAACCTGTCTAACGGACCTCATACC 895

Qy 21 ThrAlaLeuMetAlaSerProGluLysArgLeuThrLeuAlaGlnValtyrglutpMet 4.0
Db 896 AACGCATCGAGCTGGCGAGAACGGCTCACGGTGTCCAGACTGAGTGGATG 955

Qy 41 ValGlnAsnValProTyrrheArgAspLysGlyAspSerAsnSerAlaGlyTrpLys 6.0
Db 1016 AATCAATTGCTCATATCTGCCCCATACAGCAAGTGTGAGATGAGGA 1075

Qy 81 AlaGlyLysSerSerTrpTrpVal1leAsnProAsp -- -AlaLysProGlyMetAsnPro 99
Db 1076 ACTGGAAAAGTTCTGTGGATGCTTAATCCAGACGGTGGCAAAGGGGGAAATCTCCT 1135

RESULT 6
US-10-341-434-52
; Sequence 52, Application US/10341434
; Publication No. US2003021583A1
; GENERAL INFORMATION:
; APPLICANT: Origene Technologies

Qy 1 LysLysThrThrArgAsnAlaTrpGlyAsnMetSerTyraGluLeuIleThr 20
Db 1366 AGGAATGTTGCTGGCCCTACTTCAGGATAAGGGACAGCAAGCTGGGAAACCTGTCCATACGGCGACCTGATCAC 1425

Qy 21 ThrAlaLeuMetAlaSerProGluLysArgLeuThrLeuAlaGlnValtyrglutpMet 4.0
Db 1426 CGGCCATCGAGCTGGCTACTCTGCCAGAACGGCTACTGTCCAGATACCGAGTGGATG 1485

Qy 41 ValGlnAsnValProTyrrheArgAspLysGlyAspSerAsnSerAlaGlyTrpLys 6.0
Db 1486 GTGGTTGGCTACTTCAGGATAAGGGACAGCAAGCTGGCGCTGAAAG 1545

Qy 61 AsnSerIleArgHisAsnLeuSerLeuHisSerArgPheMetArgIleGlnAsnGluGly 80
Db 1546 AACCCATCGGCAACACTGTCAGTCGATTAGTCAGATGAGGAA 1605

Qy 81 AlaGlyLysSerSerTrpTrpVal1leAsnProAsp -- -AlaLysProGlyMetAsnPro 99
Db 1606 ACTGGCAGAGCTCTTGGATCATCAACCCCTGATGGGAAAGGGAAAAGGCC 1665

; TITLE OF INVENTION: Differentially Regulated Prostate Cancer Genes
 ; FILE REFERENCE: 9U 204 205 R1
 ; CURRENT APPLICATION NUMBER: US10/341,434
 ; CURRENT FILING DATE: 2003-07-18
 ; PRIORITY NUMBER: US 60/348,164
 ; PRIORITY FILING DATE: 2002-01-15
 ; PRIORITY APPLICATION NUMBER: US 60/348,119
 ; PRIORITY FILING DATE: 2002-01-15
 ; NUMBER OF SEQ ID NOS: 238
 ; SOFTWARE: Patentin version 3.1
 ; SEQ ID NO: 52
 ; LENGTH: 5723
 ; TYPE: DNA
 ; ORGANISM: Homo sapiens
 ; FEATURE:
 ; NAME/KEY: CDS
 ; LOCATION: (386) ... (2353)
 ;
 ; Alignment Scores:
 ; Pred. No.: 5.96e-51 Length: 5723
 ; Score: 427.50 Matches: 75
 ; Percent Similarity: 91.18% Conservative: 18
 ; Best Local Similarity: 73.53% Mismatches: 8
 ; Query Match: 77.03% Indels: 1
 ; DB: 17 Gaps: 1
 ;
 US-09-844-353A-54 (1-103) × US-10-671-074-4 (1-5723)
 ;
 ; Alignment Scores:
 ; Pred. No.: 1 LysLysThrThrArgAsnAlaTrpGlyAsnMetSerTyrAlaGluLeuIleThr 20
 ; Score: 836 AAGAGAGCTGGTCGCCGAACCGTGTGGCAACCTGCTCATCGCTGAGATCTAGGATG 895
 ; Percent Similarity: 91.18% Matches: 75
 ; Best Local Similarity: 73.53% Conservative: 18
 ; Query Match: 77.03% Mismatches: 8
 ; DB: 16 ValGlnAsnValProTyrrPheArgAspLysGLYAspSerAsnSerAlaGlyTrpLys 60
 ;
 US-09-844-353A-54 (1-103) × US-10-341-434-52 (1-5723)
 ;
 ; Alignment Scores:
 ; Pred. No.: 1 LysLysThrThrArgAsnAlaTrpGlyAsnMetSerTyrAlaGluLeuIleThr 20
 ; Score: 836 AAGAGAGCTGGTCGCCGAACCGTGTGGCAACCTGCTCATCGCTGAGATCTAGGATG 895
 ; Percent Similarity: 91.18% Matches: 75
 ; Best Local Similarity: 73.53% Conservative: 18
 ; Query Match: 77.03% Mismatches: 8
 ; DB: 16 ValGlnAsnValProTyrrPheArgAspLysGLYAspSerAsnSerAlaGlyTrpLys 60
 ;
 US-09-844-353A-54 (1-103) × US-10-341-434-52 (1-5723)
 ;
 ; Alignment Scores:
 ; Pred. No.: 1 LysLysThrThrArgAsnAlaTrpGlyAsnMetSerTyrAlaGluLeuIleThr 20
 ; Score: 836 AAGAGAGCTGGTCGCCGAACCGTGTGGCAACCTGCTCATCGCTGAGATCTAGGATG 895
 ; Percent Similarity: 91.18% Matches: 75
 ; Best Local Similarity: 73.53% Conservative: 18
 ; Query Match: 77.03% Mismatches: 8
 ; DB: 16 ValGlnAsnValProTyrrPheArgAspLysGLYAspSerAsnSerAlaGlyTrpLys 60
 ;
 US-09-844-353A-54 (1-103) × US-10-341-434-52 (1-5723)
 ;
 ; Alignment Scores:
 ; Pred. No.: 1 LysLysThrThrArgAsnAlaTrpGlyAsnMetSerTyrAlaGluLeuIleThr 20
 ; Score: 836 AAGAGAGCTGGTCGCCGAACCGTGTGGCAACCTGCTCATCGCTGAGATCTAGGATG 895
 ; Percent Similarity: 91.18% Matches: 75
 ; Best Local Similarity: 73.53% Conservative: 18
 ; Query Match: 77.03% Mismatches: 8
 ; DB: 16 ValGlnAsnValProTyrrPheArgAspLysGLYAspSerAsnSerAlaGlyTrpLys 60
 ;
 US-09-844-353A-54 (1-103) × US-10-341-434-52 (1-5723)
 ;
 ; Alignment Scores:
 ; Pred. No.: 1 LysLysThrThrArgAsnAlaTrpGlyAsnMetSerTyrAlaGluLeuIleThr 20
 ; Score: 836 AAGAGAGCTGGTCGCCGAACCGTGTGGCAACCTGCTCATCGCTGAGATCTAGGATG 895
 ; Percent Similarity: 91.18% Matches: 75
 ; Best Local Similarity: 73.53% Conservative: 18
 ; Query Match: 77.03% Mismatches: 8
 ; DB: 16 ValGlnAsnValProTyrrPheArgAspLysGLYAspSerAsnSerAlaGlyTrpLys 60
 ;
 US-09-844-353A-54 (1-103) × US-10-341-434-52 (1-5723)
 ;
 ; Alignment Scores:
 ; Pred. No.: 1 LysLysThrThrArgAsnAlaTrpGlyAsnMetSerTyrAlaGluLeuIleThr 20
 ; Score: 836 AAGAGAGCTGGTCGCCGAACCGTGTGGCAACCTGCTCATCGCTGAGATCTAGGATG 895
 ; Percent Similarity: 91.18% Matches: 75
 ; Best Local Similarity: 73.53% Conservative: 18
 ; Query Match: 77.03% Mismatches: 8
 ; DB: 16 ValGlnAsnValProTyrrPheArgAspLysGLYAspSerAsnSerAlaGlyTrpLys 60
 ;
 RESULT 8
 ; Sequence 11, Application US/10671074
 ; Publication No. US2004009745A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Dobie, Kenneth W.
 ; APPLICANT: Bhanot, Sanjay
 ; APPLICANT: Veniant-Ellison, Murielle
 ; APPLICANT: Lindberg, Richard A.
 ; APPLICANT: Shutter, John R.
 ; TITLE OF INVENTION: MODULATION OF FORKHEAD BOX O1A EXPRESSION
 ; FILE REFERENCE: AMGN0001-101
 ; CURRENT APPLICATION NUMBER: US/10/671,074
 ; CURRENT FILING DATE: 2003-09-25
 ; PRIOR APPLICATION NUMBER: US 10/260,203
 ; NUMBER OF SEQ ID NOS: 176
 ; SEQ ID NO 11 LENGTH: 4945
 ; TYPE: DNA
 ; ORGANISM: M. musculus
 ; FEATURE:
 ; NAME/KEY: CDS
 ; LOCATION: (429) ... (2387)
 ;
 RESULT 7
 US-10-671-074-4
 ; Sequence 4, Application US/10671074
 ; Publication No. US2004009745A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Dobie, Kenneth W.
 ; APPLICANT: Bhanot, Sanjay
 ; APPLICANT: Veniant-Ellison, Murielle
 ; APPLICANT: Lindberg, Richard A.
 ; APPLICANT: Shutter, John R.
 ; TITLE OF INVENTION: MODULATION OF FORKHEAD BOX O1A EXPRESSION
 ; FILE REFERENCE: AMGN0001-101
 ; CURRENT APPLICATION NUMBER: US/10/671,074
 ; CURRENT FILING DATE: 2003-09-25
 ; PRIOR APPLICATION NUMBER: US 10/260,203
 ; NUMBER OF SEQ ID NOS: 176
 ; SEQ ID NO 4 LENGTH: 5723
 ;
 Alignment Scores:
 ; Pred. No.: 6.79e-51 Length: 4945
 ; Score: 426.50 Matches: 75
 ; Percent Similarity: 91.18% Conservative: 18

Best Local Similarity: 73.53%
 Query Match: 76.85%
 DB: 17 Gaps: 1

US-09-844-353A-54 (1-103) × US-10-671-074-11 (1-4945)

Qy 1 LysLysThrThrArgArgAsnAlaTrrGlyAsnMetSerTyrAlaGluLeuIleThr 20
 Db 870 NAGCCAGTCGCGCGAACGGTGGCAACCTCGACGCCAGTCAAC 929

Qy 21 ThralAlaMetAlaSerProGluLysArgLeuThrIleAlaGlnValNaltryrGluTrpMet 40
 Db 775 CAGCCACTGAAGCGCCCGGAAAGCACTGACACTGCCCAGATCTAGAGTGGATG 834

Qy 41 ValGlnAsnValProTyPheAArgAspLysGlyAspSerAsnSerTyrAlaGlyTrpLys 60
 Db 835 GTCGGTACTGTACCTACTCTAGGAGGTGACAAGCTGAGGTTGACAA 894

Qy 61 AsnSerIleArgHisAsnLeuSerTyrAlaGluGlyLeGlnAngGluGly 80
 Db 895 AACTGATCGCCAAACCTGTCCTGCACGAAAGTTCACAGTTCACACGAGGCC 954

Qy 81 AlaLysLysSerTrrTrpValIleAsnProAsp---AlaLysProGlyMetAsnPro 99
 Db 955 ACCCGAAARAGCTTGTTGGATGCTGAACTTGAGGAGGCAAGGGCAAGGCC 1014

Qy 100 ArgArg 101
 Db 1015 CGCCGC 1020

RESULT 10
 US-10-037-270-314 ; Sequence 314, Application US/10037270
 ; Publication No. US20030104529A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Tang, Y. Tom
 ; APPLICANT: Liu, Chenghua
 ; APPLICANT: Asundi, Vinod
 ; APPLICANT: Zhang, Jie
 ; APPLICANT: Ren, Feiyan
 ; APPLICANT: Chen, Rui-hong
 ; APPLICANT: Zhao, Qing A.
 ; APPLICANT: Wehrman, Tom
 ; APPLICANT: Xue, Aidong J.
 ; APPLICANT: Yang, Yonghong
 ; APPLICANT: Wang, Jian-Rui
 ; APPLICANT: Zhou, Ping
 ; APPLICANT: Ma, Yanqing
 ; APPLICANT: Wang, Dunrui
 ; APPLICANT: Wang, Zhiwei
 ; APPLICANT: Tillinghast, John
 ; APPLICANT: Demanac, Radoe T.
 ; TITLE OF INVENTION: No. US20030104529A1 Nucleic Acids and
 ; TITLE OF INVENTION: Polypeptides
 ; FILE REFERENCE: 7841P2B
 ; CURRENT APPLICATION NUMBER: US/10/037-270
 ; CURRENT FILING DATE: 2002-01-04
 ; PRIOR APPLICATION NUMBER: US20030104529A1
 ; PRIOR FILING DATE: 2000-04-25
 ; PRIOR APPLICATION NUMBER: 09/1488,725
 ; PRIOR FILING DATE: 2000-01-21
 ; NUMBER OF SEQ ID: NOS: 1104
 ; SOFTWARE: pt_FL_genes Version 1.0
 ; LENGTH: 3394
 ; SEQ ID NO: 314
 ; TYPE: DNA
 ; ORGANISM: Homo sapiens
 ; FEATURE: CDS
 ; NAME/KEY: CDS
 ; LOCATION: (417)...(1934)
 ; LENGTH: 3394
 ; Gaps: 1

RESULT 9
 US-10-261-517-22 ; Sequence 22, Application US/10261517
 ; Publication No. US20030143583A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Climent-Johansson, Isabel
 ; APPLICANT: Dahlman-Wright, Karin
 ; APPLICANT: Lake, Staffan
 ; APPLICANT: Wasserman, Wyeth
 ; TITLE OF INVENTION: NOVEL RESPONSE ELEMENT
 ; FILE REFERENCE: 13445-032001
 ; CURRENT APPLICATION NUMBER: US/10/261,517
 ; CURRENT FILING DATE: 2002-10-01
 ; PRIOR APPLICATION NUMBER: US/09/645,629
 ; PRIOR FILING DATE: 2000-08-24
 ; PRIOR APPLICATION NUMBER: US 60/151,867
 ; PRIOR FILING DATE: 1999-08-31
 ; PRIOR APPLICATION NUMBER: SE 9903009-0
 ; PRIOR FILING DATE: 1999-08-26
 ; SOFTWARE: FastSEQ for Windows Version 4.0
 ; SEQ ID NO: 22
 ; LENGTH: 1073
 ; TYPE: DNA
 ; ORGANISM: Artificial Sequence
 ; FEATURE: OTHER INFORMATION: synthetically generated GST-AfxxBD construct
 US-10-261-517-22

Alignment Scores:
 Pred No.: 1.72e-49 Length: 1073
 Score: 410.50 Matches: 73
 Percent Similarity: 86.27% Conservative: 15
 Best Local Similarity: 71.57% Misnatches: 13
 Query Match: 73.56% Indels: 1
 Gaps: 1

DB: 715 CGGAGGGGGGTCGGCGGAATGCCCTGGAAATCATGAGAACTCATGAC 774

Qy 1 LysLysThrThrArgArgAsnAlaTrrGlyAsnMetSerTyrAlaGluLeuIleThr 20
 Db 715 CGGAGGGGGGTCGGCGGAATGCCCTGGAAATCATGAGAACTCATGAC 774

Qy 1 LysLysThrThrArgArgAsnAlaTrrGlyAsnMetSerTyrAlaGluLeuIleThr 20
 US-09-844-353A-54 (1-103) × US-10-037-270-314 (1-3394)

RESULT 11
US-10-117-722-314
; Sequence 314, Application US/10117722
; GENERAL INFORMATION:
; Publication No. US2003021974A1
; APPLICANT: Tang, Y., Tom
; APPLICANT: Liu, Chenghua
; APPLICANT: Aundi, Vinod
; APPLICANT: Zhang, Jie
; APPLICANT: Drmanac, Radoje T.
; TITLE OF INVENTION: NO-US2003021974A1el Nucleic Acids and
; FILE REFERENCE: 784C1P2BCIP
; CURRENT APPLICATION NUMBER: US/10/117,722
; CURRENT FILING DATE: 2002-04-04
; PRIOR APPLICATION NUMBER: 09/620,312
; PRIOR FILING DATE: 2000-07-19
; PRIOR APPLICATION NUMBER: 09/552,317
; PRIOR FILING DATE: 2000-04-25
; PRIOR APPLICATION NUMBER: 09/488,725
; PRIOR FILING DATE: 2000-01-21
; NUMBER OF SEQ ID NOS: 1104
; SOFTWARE: pt_FL_genes Version 1.0
; SEQ ID NO: 314
; LENGTH: 3394
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (417) .. (1934)
; US-10-117-722-314

Alignment Scores:
Pred. No.: 8.6e-49 Length: 3394
Score: 410.50 Conservative: 73
Percent Similarity: 86.27% Mismatches: 13
Best Local Similarity: 71.57% Indels: 1
Query Match: 73.96% Gaps: 1
DB: 16

US-09-844-353A-54 (1-103) x US-10-117-722-314 (1-3394)

Qy 1 LysLysThrThrArgAsnAlaTrpGlyAsnMetSerTyraGluLeuIleThr 20
Db 690 CGCAAGCCGGCTCCCCGGGAATGCTGGAAATCATGAGAACTCATAGC 749

Qy 21 ThialaileMetalaserProGlyLysArgLeuIleValtrycLutrPmet 40
Db 750 CAGGCCATTGAAAGCCCGAGAACGACTTGCCAGATCTACAGGGATG 809

Qy 41 ValGlnAsnValProtYrPheArgAspLysGlyAspSerAsnSerAlaGlyTIPlys 60
Db 810 GTCCGTACTGAACTCTACTGAGAACAAAGGGTGCAGAACGAGGAAAG 869

Qy 61 AsnSerIleArgHisAsnLeuSerLeuHisSerArgPheMetArgIleGlnAsnGluGly 80
Db 870 AACTCATCCCAACCTGCTCCAGRCGCAACTTCATCAAGTTCAACAGGGCC 929

Qy 81 AlaGlyLysSerSetTrpTrpValleAsnProAsp---AlaLysProGlyMetAsnPro 99
Db 930 ACCGGAAAAGCTCTGGGATGCTGAGGAGCAAGGGCAAAAGCCCC 989

Qy 100 ArgArg 101
Db 990 CGCCGC 995

RESULT 12
US-09-205-658-44
; Sequence 44, Application US/09205658
; GENERAL INFORMATION:
; APPLICANT: Ruvkin, Gary
; APPLICANT: Ogg, Scott
; TITLE OF INVENTION: THERAPEUTIC AND DIAGNOSTIC TOOLS FOR
; IMPAIRED GLUCOSE TOLERANCE CONDITIONS
; CURRENT APPLICATION NUMBER: US/09/205,658
; CURRENT FILING DATE: 1998-12-03
; EARLIER APPLICATION NUMBER: 08/857,076
; EARLIER FILING DATE: 1997-05-15
; EARLIER APPLICATION NUMBER: 08/888,534
; EARLIER FILING DATE: 1997-07-07
; EARLIER APPLICATION NUMBER: 00786/351004
; FILE REFERENCE: 00786/351004
; NUMBER OF SEQ ID NOS: 328
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO: 44
; LENGTH: 2704
; TYPE: DNA
; ORGANISM: Caenorhabditis elegans
US-09-205-658-44

Alignment Scores:
Pred. No.: 2.87e-44 Length: 2704
Score: 378.50 Matches: 67
Percent Similarity: 75.93% Conservative: 15
Best Local Similarity: 62.04% Mismatches: 21
Query Match: 68.20% Indels: 5
DB: 9 Gaps: 1

US-09-844-353A-54 (1-103) x US-09-205-658-44 (1-2704)

Qy 1 LysLysThrThrArgAsnAlaTrpGlyAsnMetSerTyrr 15
Db 398 AGAACCCAACCGATCAATTGGCACAGAAACCGAAATCCATGGGTGAGAATCTAT 457

Qy 16 AlaGluLeuIleThrThrAlaIleMetAlaSerProGlyLysArgLeuThrLeuAlaGln 35
Db 458 TCGGATCATGCCCAGATGGATGCGGCCAGGGTTAAACTCAATGAG 517

Qy 36 ValTyrClutPmetValGlnAsnvalProtYrPheArgAspLysGlyAspSerAsnSer 55
Db 518 ATTACATGCTCTGATAATCCCTACTTGGAGAACGACTGTCGGAG 577

Qy 56 SerAlaGlyTrpLysAsnSerIleArgHisAsnLeuSerLeuHisSerArgPheMetArg 75
Db 578 GCGCCGGATGAAAGACTCGATCCGTCACATCTGTCCTCATTCATGTTCATGCGA 637

Qy 76 IleGlnAsnGluGlyLysSerSerTrpValleAsnProAspAlaLysPro 95
Db 638 ATTCAGATGAGGAGSCGGAAAAGCTCTGGGTTATAATCCAGATGCAAAGCCA 697

Qy 96 GlyMetAsnProArgArgThrArg 103

RESULT 13
Db 698 GGAATGAACTCACGGCGTACCGT 721
; TITLE OF INVENTION: IMPAIRED GLUCOSE TOLERANCE CONDITIONS
; FILE REFERENCE: 00786/35104
; CURRENT FILING DATE: 2001-09-25
; PRIORITY APPLICATION NUMBER: US/09/963,693
; PRIORITY FILING DATE: 1997-07-07
; PRIORITY APPLICATION NUMBER: 08/057,076
; PRIORITY FILING DATE: 1997-05-15
; PRIORITY APPLICATION NUMBER: 08/788,534
; PRIORITY FILING DATE: 1997-07-07
; PRIORITY APPLICATION NUMBER: US98/10080
; PRIORITY FILING DATE: 1998-12-03
; NUMBER OF SEQ ID NOS: 328
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 44
; LENGTH: 2704
; TYPE: DNA
; ORGANISM: Caenorhabditis elegans
US-09-963-693-44

Alignment Scores:
Pred. No.: 2.87e-44 Length: 2704
Score: 378.50 Matches: 67
Percent Similarity: 75.3% Conservative: 15
Best Local Similarity: 62.04% Mismatches: 21
Query Match: 68.20% Indels: 5
DB: 10 Gaps: 1

US-09-844-353A-54 (1-103) x US-09-963-693-44 (1-2704)

Qy 1 LysLysThrThr-----ThrArgArgAsnAlaTrpGlyAsnMetSerThr 15
Db 398 AAGAGCCACCATCAATTGGCACAGAGAACCGAAATCCATGGTGGAGGATCTPAT 457

Alignment Scores:
Pred. No.: 2.87e-44 Length: 2704
Score: 378.50 Matches: 67
Percent Similarity: 75.3% Conservative: 15
Best Local Similarity: 62.04% Mismatches: 21
Query Match: 68.20% Indels: 5
DB: 9 Gaps: 1

Qy 36 ValTyRGlutpMetValGlnAsnValpTyrpheArgAspLysGlyAspSerAsnSer 55
Db 518 ATTATCATGGTTCTCTGATAATATTCCTACTTGTGAGAAAGATCTAGTCCGAGAG 577

Qy 56 SerAlaGlyTrpLysAsnSerIleArgHisAsnLeuSerLeuHisSerArgPheMetArg 75
Db 458 TCGGATATACTGGCAAACCATGGAACTGGGCCAGAGGGCTTAACCTCAATGAG 517

Qy 76 IleGlnAsnGluGlyAlaGlyLysSerSerTrpValIleAsnProAspAlaLysPro 95
Db 638 ATTCAAGATGAAAGGAGCCGAAAGCTGTTATTAACTCCAGATGAAAGCCA 697

RESULT 15
US-09-186-839-1
; Sequence 1, Application US/10186839
; Publication No. US20030096957A1
; GENERAL INFORMATION:
; APPLICANT: Bowen, Benjamin R.
; APPLICANT: Whelan, James P.
; APPLICANT: Yang, Zhenyu
; TITLE OF INVENTION: ARXZeta Transcription Factor Splice Form
; FILE REFERENCE: 4-3206A/USN
; CURRENT FILING DATE: 2002-10-02
; PRIORITY APPLICATION NUMBER: 60/302,134
; PRIORITY FILING DATE: 2001-06-29
; NUMBER OF SEQ ID NOS: 8
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 1
; LENGTH: 1353
; TYPE: DNA

RESULT 14
US-09-963-693-44
; Sequence 44, Application US/09963693
; Publication No. US20030181364A1
; GENERAL INFORMATION:
; APPLICANT: Ruvkin, Gary
; APPLICANT: Ogg, Scott
; TITLE OF INVENTION: THERAPEUTIC AND DIAGNOSTIC TOOLS FOR

RESULT 13
US-09-844-353a-44
; Sequence 44, Application US/09844353A
; Patent No. US20030037585A1
; GENERAL INFORMATION:
; APPLICANT: Ruvkin, Gary
; Kimura, Koutarou
; Patterson, Garth
; Ogg, Scott
; Paradis, Suzanne
; Tissenbaum, Heidi
; Morris, Jason
; Kowek, Allison
; TITLE OF INVENTION: THERAPEUTIC AND DIAGNOSTIC TOOLS FOR
; IMPAIRED GLUCOSE TOLERANCE CONDITIONS
; FILE REFERENCE: 00786/35105
; CURRENT FILING DATE: 2001-04-27
; PRIORITY APPLICATION NUMBER: US/08/857,076
; PRIORITY FILING DATE: 1997-05-15
; NUMBER OF SEQ ID NOS: 114
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 44
; LENGTH: 2704
; TYPE: DNA
; ORGANISM: Caenorhabditis elegans
US-09-844-353A-44

Alignment Scores:
Pred. No.: 2.87e-44 Length: 2704
Score: 378.50 Matches: 67
Percent Similarity: 75.3% Conservative: 15
Best Local Similarity: 62.04% Mismatches: 21
Query Match: 68.20% Indels: 5
DB: 9 Gaps: 1

US-09-844-353A-54 (1-103) x US-09-844-353A-44 (1-2704)

Qy 1 LysLysThrThr-----ThrArgArgAsnAlaTrpGlyAsnMetSerThr 15
Db 398 AAGAGCCACCATCAATTGGCACAGAGAACCGAAATCCATGGTGGAGGATCTPAT 457

Qy 16 AlaGluLeuIleThrThrAlaLeuMetAlaSerProGluLysArgLeuThrLeuAlaGln 35
Db 458 TCGGATATACTGGCAAACCATGGAACTGGGCCAGAGGGCTTAACCTCAATGAG 517

Qy 578 GCGCCGATGGAGAACTGATCCTGAACTCTGCTTCATGGGA 637

Qy 76 IleGlnAsnGluGlyAlaGlyLysSerSerTrpValIleAsnProAspAlaLysPro 95
Db 638 ATTCAAGATGAAAGGAGCCGAAAGCTGTTATTAACTCCAGATGAAAGCCA 697

Qy 96 GlyMetAsnProArgArgGlyArg 103
Db 698 GGATGATCCACGGCGTACACST 721

RESULT 15
US-09-186-839-1
; Sequence 1, Application US/10186839
; Publication No. US20030096957A1
; GENERAL INFORMATION:
; APPLICANT: Bowen, Benjamin R.
; APPLICANT: Whelan, James P.
; APPLICANT: Yang, Zhenyu
; TITLE OF INVENTION: ARXZeta Transcription Factor Splice Form
; FILE REFERENCE: 4-3206A/USN
; CURRENT FILING DATE: 2002-10-02
; PRIORITY APPLICATION NUMBER: 60/302,134
; PRIORITY FILING DATE: 2001-06-29
; NUMBER OF SEQ ID NOS: 8
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 1
; LENGTH: 1353
; TYPE: DNA

ORGANISM: Homo Sapien
US-10-186-839-1

Search completed: July 29, 2004, 13:22:45
Job time : 364 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using SW model

Run on: July 23, 2004, 10:21:40 ; Search time 39 Seconds
(without alignments)
833.292 Million cell updates/sec

Title: US-09-844-353A-54

Perfect score: 555

Sequence: 1 KRTTRRNAGNMSYELIT.....SSWWVTPDAKPGMNPRTTR 103

Scoring table: BLASTM62

Gapop: 10.0 , Gapext 0.5

Searched: 1017041 seqs, 315518202 residues

Total number of hits satisfying chosen parameters: 1017041

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : SPTRNMBL_25;*

- 1: sp_archea;*
- 2: sp_bacteria;*
- 3: sp_fungi;*
- 4: sp_human;*
- 5: sp_invertebrate;*
- 6: sp_mammal;*
- 7: sp_mhc;*
- 8: sp_organelle;*
- 9: sp_phage;*
- 10: sp_plant;*
- 11: sp_rabbit;*
- 12: sp_virus;*
- 13: sp_vertebrate;*
- 14: sp_unclassified;*
- 15: sp_rvirus;*
- 16: sp_bacteri;*
- 17: sp_archaea;*

Q9vfn8 drosophila
Q8bar3 mus musculu
Q17593 caenorhabdi
Q22510 caenorhabdi
Q9vtf7 drosophila
Q8it15 nematopsis
Q8jif5 xenopus lae
Q9tzkl caenorhabdi
Q7ybz3 saccoGLOSSA
Q7zyq0 xenopus lae
Q7tico brachydanio
Q86lt7 branchiosteo
Q78x57 avian sroco
Q61733 branchiosteo
Q9ae25 brachydanio
Q92215 mus musculu
Q9d299 mus musculu
Q8sz55 drosophila
Q8ovp3 mus musculu
Q800k3 oryzias lat
Q86xt7 homo sapiens
Q92215 mus musculu
Q9yhc5 drosophila
Q8ovp3 mus musculu
Q800k3 oryzias lat
Q86xt7 homo sapiens
Q92215 mus musculu
Q9d299 mus musculu
Q8sz55 drosophila
Q73785 brachydanio
Q9vhb2 xenopus lae
Q9pvz3 xenopus lae
Q3440 gallus gallus
Q7t2g3 brachydanio
Q8bik9 mus musculu

ALIGNMENTS

RESULT 1

ID	QBXX10	PRELIMINARY;	PRT;	407 AA.
AC	QBXX10;			
DT	01-OCT-2002 (TREMBLrel.	22,	Created)	
DT	01-OCT-2002 (TREMBLrel.	22,	Last sequence update)	
DT	01-JUN-2003 (TREMBLrel.	24,	Last annotation update)	
DE	DAF16	protein (corresponding sequence R13H8.id).		
GN	R13H8.1 OR DAF-16.			
OS	Caenorhabditis elegans			
OC	Bukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditoidae;			
OC	Rhabditidae; Peioderinae; Caenorhabditis.			
RN	[1] NCBI_TaxID=6239;			
RN	SEQUENCE FROM N.A.			
RC	STRAIN-Bristol N2;			
RX	Medline-99059613; PubMed-9851916;			
RA	Waterson R.;			
RT	"Genome sequence of the nematode C. elegans: a platform for investigating biology." The C. elegans Sequencing Consortium. ;			
RL	Science 282:2012-2018(1998).			
RN	[2]			
RP	SEQUENCE FROM N.A.			
RC	STRAIN-Bristol N2;			
RA	Jones K., Hinds K., Sutterer C., Coffman M. ;			
RL	"The sequence of C. elegans contid R13H8.1;"			
RL	Submitted (JAN-1998) to the EMBL/GenBank/DDBJ databases.			
RN	[3]			
RP	SEQUENCE FROM N.A.			
RC	STRAIN-Bristol N2;			
RA	Waterson R.;			
RL	Submitted (JUN-2002) to the EMBL/GenBank/DDBJ databases.			
DR	EMBL; AF039177; AAM51498.1; -.			
DR	Wormpep; R13H8.1; -.			
DR	GO; GO-0005634; C:nucleus; IEA.			
DR	GO; GO-0003030; F:regulation of transcription factor activity; IEA.			
DR	InterPro; IPR01766; TF_Fork_head; 1.			
DR	PFam; PF00250; Fork_head; 1.			

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the distribution and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query	Match	Length	DB	ID	Description
1	54.9	98.9	407	5	Q8mx10	Q8mx10 caenorhabdi	
2	54.9	98.9	508	5	O16849	O16849 caenorhabdi	
3	54.9	98.9	510	5	O18676	O18676 caenorhabdi	
4	45.0	81.1	613	5	Q95v55	Q95v55 drosophila	
5	44.0	79.4	672	11	Q9WTH4	Q9WTH4 mus musculu	
6	43.5	78.5	727	13	Q9WTF9	Q9WTF9 gallus gallus	
7	43.3	78.1	664	13	Q9YK2	Q9YK2 xiphophorus	
8	42.7	77.0	653	11	Q810ws	Q810ws spermophilus	
9	42.7	77.0	662	11	Q8M69	Q8M69 sus scrofa	
10	42.6	76.8	652	11	Q9UW4	Q9UW4 mus musculu	
11	42.6	76.8	652	11	Q9WTH5	Q9WTH5 mus musculu	
12	42.5	76.7	651	13	Q9WTF8	Q9WTF8 brachydanio	
13	37.2	67.1	530	5	O16550	O16550 caenorhabdi	
14	33.9	61.2	333	11	Q9SS26	Q9SS26 rattus norvegicus	
15	28.6	51.5	204	13	Q9PVK6	Q9PVK6 brachydanio	
16	26.7	48.2	593	5	Q8TS12	Q8TS12 anopheles g	

DR	PRINTS;	PRO0053; FORKHEAD.		DR	PRINTS; PRO0053; FORKHEAD.	
DR	ProDom;	PD000425; TF_Fork_head; 1.		DR	ProDom; PD000425; TF_Fork_head; 1.	
DR	SMART;	SM00339; FH; <u>1</u> .		DR	SMART; SM00339; FH; <u>1</u> .	
DR	PROSITE;	PS00658; FORK_HEAD_2; 1.		DR	PROSITE; PS00658; FORK_HEAD_2; 1.	
DR	PROSITE;	PS50039; FORK_HEAD_3; 1.		DR	PROSITE; PS50039; FORK_HEAD_3; 1.	
SEQ	SEQUENCE	407 AA; 44727 MW; 4E5CA10AF689DBEF CRC64;		SEQ	SEQUENCE 508 AA; 55577 MW; 2C0CF97657CD6350 CRC64;	
Query Match	98 %;	Score 549; DB 5; Length 407;		Query Match	98.9 %;	Score 549; DB 5; Length 508;
Best Local Similarity	99.0 %;	Pred. No. 7.1e-54;		Best Local Similarity	99.0 %;	Pred. No. 9.4e-54;
Matches	102;	Conservative 0; Mismatches 0;		Matches	102;	Conservative 0; Mismatches 0;
QY	1	KKTTTIRNANGNMSTAELITTAIMASPEKRITLAQQYENWVQNPYFRDGDSNSAGWK 60		QY	1 KKTTTIRNANGNMSTAELITTAIMASPEKRITLAQQYENWVQNPYFRDGDSNSAGWK 60	
Db	32	KKTTTIRNANGNMSTAELITTAIMASPEKRITLAQQYENWVQNPYFRDGDSNSAGWK 91		Db	133 KKTTTIRNANGNMSTAELITTAIMASPEKRITLAQQYENWVQNPYFRDGDSNSAGWK 192	
QY	61	NSIRHNLSSLHFRMFLIQNEGAGKSSMWVINVDAKPGMNPRTR 103		QY	61 NSIRHNLSSLHFRMFLIQNEGAGKSSMWVINVDAKPGMNPRTR 103	
Db	92	NSIRHNLSSLHFRMFLIQNEGAGKSSMWVINVDAKPGRNPRTR 134		Db	193 NSIRHNLSSLHFRMFLIQNEGAGKSSMWVINVDAKPGRNPRTR 235	
RESULT 2						
O16849	ID	O16849 PRELIMINARY; PRT; 508 AA.		O16876	ID	O16876 PRELIMINARY; PRT; 510 AA.
AC	AC	O16849;		AC	AC O16876;	
DT	DT	01-JAN-1998 (TREMBLrel. 05, Created)		DT	DT 01-JAN-1998 (TREMBLrel. 05, Created)	
DT	DT	01-JAN-1998 (TREMBLrel. 05, Last sequence update)		DT	DT 01-JAN-1998 (TREMBLrel. 05, Last sequence update)	
DT	DT	01-JUN-2003 (TREMBLrel. 24, Last annotation update)		DT	DT 01-JUN-2003 (TREMBLrel. 24, Last annotation update)	
DB	DB	Fork head-related transcription factor DAF-16a2 (Hypothetical protein R13H8.1b).		DB	DB DAF-16 (Hypothetical protein R13H8.1C).	
DE	DE	R13H8.1b).		GN	GN DAF-16 OR R13H8.1.	
DE	DE	DAF-16 OR R13H8.1.		OS	OS Caenorhabditis elegans.	
OS	OS	Caenorhabditis elegans.		OC	OC Buxaryota; Metazoa; Chromadorea; Rhabditida; Rhabditoidea;	
OC	OC	Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditoidea;		OC	OC Caenorhabditis; Peloderaiae; Caenorhabditis.	
OX	OX	NCBI TAXID=6239;		OX	OX NCBI TAXID=6239;	
[1]	[1]	SEQUENCE FROM N.A.		RN	RN [1]	
RP	RP	SEQUENCE FROM N.A.		RP	RP SEQUENCE FROM N.A.	
RC	RC	STRAIN=N2;		RC	RC STRAIN=BRISTOL N2;	
RC	RC	MEIDLINE=98013175; PubMed=9353126;		RX	RX MEDLINE=98013175; PubMed=9353126;	
RA	RA	OGG S., Paradis S., Gottlieb S., Patterson G.I., Lee L., Tissenbaum H.A., Ruvkun G.;		RA	RA Lin K., Domman J.B., Rodan A., Kenyon C.;	
RA	RA	"The Fork head transcription factor DAF-16 transduces insulin-like growth factor signaling in C. elegans.";		RA	RA "daf-16: An HNF-3/forkhead family member that can function to double the life-span of Caenorhabditis elegans.";	
RA	RA	"The Fork head transcription factor DAF-16 transduces insulin-like growth factor signaling in C. elegans.";		RT	RT Science 278:1319-1322(1997).	
RT	RT	Nature 389:994-999(1997).		RL	RL [2]	
RL	RL	SEQUENCE FROM N.A.		RN	RN SEQUENCE FROM N.A.	
RN	RN	STRAIN=Bristol N2;		RC	RC SEQUENCE FROM N.A.	
RC	RC	MEIDLINE=99069613; PubMed=9851916;		RC	RC STRAIN=Bristol N2;	
RA	RA	None;		RA	RA MEDLINE=99069613; PubMed=9851916;	
RA	RA	"Genome sequence of the nematode C. elegans: a platform for investigating biology." The C. elegans Sequencing Consortium. Science 282:2012-2018(1998).		RA	RA None;	
RA	RA	[3]		RT	RT "Genome sequence of the nematode C. elegans: a platform for investigating biology." The C. elegans Sequencing Consortium. Science 282:2012-2018(1998).	
RN	RN	SEQUENCE FROM N.A.		RL	RL [3]	
RC	RC	STRAIN=Bristol N2;		RN	RN SEQUENCE FROM N.A.	
RA	RA	Jones K., Hinds K., Suttorer C., Cofman M.;		RC	RC SEQUENCE FROM N.A.	
RA	RA	"The sequence of C. elegans cosmid R13H8.";		RA	RA Jones K., Hinds K., Suttorer C., Cofman M.;	
RA	RA	Submitted (JAN-1998) to the EMBL/GenBank/DDBJ databases.		RA	RA "The sequence of C. elegans cosmid R13H8.";	
RL	RL	[4]		RT	RT Submitted (JAN-1998) to the EMBL/GenBank/DDBJ databases.	
RN	RN	SEQUENCE FROM N.A.		RL	RL [4]	
RC	RC	STRAIN=Bristol N2;		RN	RN SEQUENCE FROM N.A.	
RA	RA	Waterston R.;		RC	RC SEQUENCE FROM N.A.	
RA	RA	"Direct Submission.";		RA	RA Jones K., Hinds K., Suttorer C., Cofman M.;	
RA	RA	Submitted (AUG-2001) to the EMBL/GenBank/DDBJ databases.		RA	RA "The sequence of C. elegans cosmid R13H8.";	
RL	RL	[5]		RT	RT Submitted (AUG-2001) to the EMBL/GenBank/DDBJ databases.	
DR	DR	EMBL; AF020343; AAB8491.1; -.		RL	RL [5]	
DR	DR	HSSP; Q63245; 2PH.		RN	RN SEQUENCE FROM N.A.	
DR	DR	TRANSPAC; T03399; -.		RC	RC SEQUENCE FROM N.A.	
DR	DR	WormPeP; R13H8.1b; CB28772.		RA	RA Waterston R.;	
DR	DR	GO:0005634; C:nucleus; IEA.		RA	RA "Direct Submission.";	
DR	DR	GO: GO:0003700; F:transcription factor activity; IEA.		RL	RL DR Submitted (AUG-2001) to the EMBL/GenBank/DDBJ databases.	
DR	DR	GO: GO:0006355; P:regulation of transcription, DNA-dependent; IEA.		DR	DR EMBL; AF020342; AAC4703.1; -.	
DR	DR	InterPro; IPR001766; TF_Fork_head.		DR	DR EMBL; AF020342; AAC84390.1; -.	
DR	DR	Pfam; PF00250; Fork_head; 1.		DR	DR EMBL; AF03917; AAK82219.1; -.	

DR	PIR; T42255; T42255.	Query Match Score 450; DB 5; Length 613;
DR	HSSP; Q63245; 2HFF..	Best Local Similarity 77.2%; Pred. No. 2.e-42;
DR	TRANSFAC; T03398; -.	Matches 78; Conservative 16; Mismatches 7; Indels 0; Gaps 0;
DR	WormPep; R13HB.1c; CE28773.	
DR	GO; GO-0005634; C:nucleus; IEA.	Qy 1 KKTTRRNAGNMYSABLTTAIMASPEKRILTAQYEWMMVQNPYFRDKGDSNSAGWK 60
DR	GO; GO-0003700; P:transcription factor activity; IEA.	Db 86 KGNSSRNAGNMYSABLTHAIGSATDKRLTSLQYEWMMVQNPYFRDKGDSNSAGWK 145
DR	GO; GO-0006555; P:regulation of transcription, DNA-dependent; IEA.	
DR	InterPro; IPR001766; TF_Fork_head.	
DR	Pfam; PF00250; Fork_head_1.	
DR	PRINTS; PRO0053; FORKHEAD.	
DR	PRODOM; PD000425; TF_Fork_head; 1.	
DR	SMART; SM00339; FH_1	
DR	PROSITE; PS00658; FORK_HEAD_2; 1.	
DR	PROSITE; PS00039; FORK_HEAD_3; 1.	
DR	SEQUENCE; 510 AA; 55867 MW; 52574F3F979B3583 CRC64;	RESULT 5
SQ		Q9WVH4 PRELIMINARY; PRT; 672 AA.
	Query Match Score 98.9%; Pred. No. 9.5e-54; Length 510;	ID Q9WVH4 PRELIMINARY; PRT; 672 AA.
	Best Local Similarity 99.0%; Mismatches 0; Indels 0; Gaps 0;	AC Q9WVH4 PRELIMINARY; PRT; 672 AA.
	Matches 102; Conservative 0; Mismatches 1; Indels 0; Gaps 0;	DT 01-NOV-1999 (TREMBLrel. 12; Created)
		DT 01-NOV-1999 (TREMBLrel. 12; Last sequence update)
Qy	1 KKTTRRNAGNMYSABLTTAIMASPEKRILTAQYEWMMVQNPYFRDKGDSNSAGWK 60	DT 01-OCT-2003 (TREMBLrel. 25; Last annotation update)
Db	135 KKTTRRNAGNMYSABLTTAIMASPEKRILTAQYEWMMVQNPYFRDKGDSNSAGWK 194	DE Forkhead protein FKH2 (Foxhead box 03).
GN		GN FOXO3 OR FKH2.
		OS Mus musculus (Mouse).
		OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
		OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
Db	61 NSIRHNLSLHSRFMRILQNEGAKSSWWINPDAKPGNPNRTR 103	NCBI_TaxID=10090; NCBI_TaxID=10090; [1].
Qy	61 NSIRHNLSLHSRFMRILQNEGAKSSWWINPDAKPGNPNRTR 103	RN SEQUENCE FROM N.A.
Db	195 NSIRHNLSLHSRFMRILQNEGAKSSWWINPDAKPGNPNRTR 237	RX MEDLINE=21251166; PubMed=11353388;
		RA Biggs W.H. III, Caveen W.K., Arden K.C.;
		RT Identification and characterization of members of the FKHRL (FOX O)
		RL Marini. Genome 12:416-425 (2001). [2]
		RN SEQUENCE FROM N.A.
		RP STRAINC57BL/6J; TISSUE=Cerebellum;
		RC MEDLINE=22354683; PubMed=12466851;
		RA The RIKEN Genome Exploration Research Group Phase I & II Team;
		RA "Analysis of the mouse transcriptome based on functional annotation of
		RT 60,770 full-length cDNAs";
		RL Nature 20:563-573 (2002).
		DR EMBL; AF114259; AAD2107; 1.
		DR EMBL; AK047413; BAC33049.1; -.
		DR HSSP; Q63245; 2HFF.
		DR TRANSFAC; T04178; -.
		DR MGD; MGI:1890081; Foxo3.
		DR GO; GO:0016563; P:transcriptional activator activity; IDA.
		DR InterPro; IPR001766; TF_Fork_head.
		DR Pfam; PF00250; Fork_head; 1.
		DR PRINTS; PR00053; FORKHEAD.
		DR Prodrom; PD00425; TF_Fork_head; 1.
		DR GO; GO:0016563; P:transcriptional activator activity; IDA.
		DR SMART; SM00339; FH_1.
		DR PROSITE; PS00058; FORK_HEAD_2; 1.
		DR PROSITE; PS50039; FORK_HEAD_3; 1.
		DR SEQUENCE 672 AA; 71064 MW; EC21.8D9BA0C1DAC5 CRC64;
		RESULT 6
		Q9W7F9 PRELIMINARY; PRT; 727 AA.

Q9W7F9;	AC DT DT DT DT DE GN OS RN [1]	DR PFF0250; Fork_head; 1. PRINTS; PR00053; FORKHEAD. DR ProdOm; PD000425; TF Fork_head; 1. DR SMART; SM00339; FH; 1. DR PROSTE; PS00058; FORK_HEAD-2; DR PROSTE; PS50039; FORK_HEAD-3; 1. DR SEQUENCE 664 AA; 7050 MW; 268D3F3FF64D7CC0 CRC64;
Gallus gallus (Chicken).	OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Buteleostomi; Archosauria; Aves; Neognathae; Galliformes; Phasianinae; Gallus.	Query Match 78.1%; Score 433.5; DB 13; Length 664; Best Local Similarity 74.5%; Pred. No. 1.8e-40; Matches 76; Conservative 17; Mismatches 8; Indels 1; Gaps 1;
FKHR.	OC FKHR protein FKHR.	Qy 1 KKTTRRNAGNMYSAYELITTAIMASPEKRILTAQYEMVQNYFRDKGDNSAGWK 60 Db 127 RKASSRNRNAGNLSTADLTKAESSAEKRUTLSQIYDMVNTSIPFKDKGDNSAGWK 186
[1]	SEQUENCE FROM N.A. MEDLINE=21251166; PubMed=11353388; RX BIGGS W.H. III., Cavenie W.K., Arden K.C.; RT Identificaton and characterization of members of the FKHR (FOX O) subclass of winged-helix transcription factors in the mouse. " ; RT Mamm. Genome 12:416-425(2001). EMLBL; AF11261; AAD42109_1; - . HSSP; Q63245; 2HFH.	Qy 61 NSIRHNLSHSRMFLQNEGAKGKSSWWVNWPD-AKGGMNPRR 101 Db 187 NSIRHNLSHSRFIRVQNEGIGKSSWWMINPEGGKSKAPRR 228
TRANSFAC;	DR TRANSFAC; T04202; - . DR GO; GO:0005634; C:nucleus; IEA. GO; GO:0003700; F:transcription factor activity; IEA. DR GO; GO:006355; P:regulation of transcription, DNA-dependent; IEA. DR InterPro; IPR00166; TF Fork_head. DR PROSTE; PR000250; Fork_head; 1. DR SMART; SM00339; FH; 1. DR PROSTE; PS000425; FORKHEAD. DR SMART; SM00339; FH; 1. DR PROSTE; PS000658; FORK_HEAD-2; 1. DR PROSTE; PS000339; FORK_HEAD-3; 1. DR SEQUENCE 727 AA; 77290 MW; 256A2220410E075E CRC64;	RESULT 8 Q810W5 ID Q810W5 PRELIMINARY; PRT; 653 AA. AC Q810W5; Q810W5; PRELIMINARY; PRT; 653 AA. DR DT 01-JUN-2003 ("TREMBlrel. 24, Created) DT 01-JUN-2003 ("TREMBlrel. 24, Last sequence update) DT 01-OCT-2003 ("TREMBlrel. 25, Last annotation update) DE Forthead box Ota protein. FOXO1A. GN Spermophilus tridecemlineatus (Thirteen-lined ground squirrel). OS Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Sciuridae; Spermophilus. OC NCBI_TaxID=43179; RN [1] RN [1] SEQUENCE FROM N.A.
Query Match 78.5%; Score 435.5; DB 13; Length 727; Best Local Similarity 74.5%; Pred. No. 1.2e-40; Matches 76; Conservative 18; Mismatches 7; Indels 1; Gaps 1;	RC TISSUE=Brain; RA Cai D.; Hallenbeck J.M.; RT "Cloning and characterization of a forkhead transcriptional factor from thirteen-lined ground squirrel."; RL Submitted (MAR 2003) to the EMBL/GenBank/DBJ databases. DR EMBL; AY25525; AA072710_1; - . DR GO; GO:0005634; C:nucleus; IEA. DR GO; GO:0006370; P:transcription factor activity; IEA. DR GO; GO:0006355; P:regulation of transcription, DNA-dependent; IEA. DR InterPro; IPR001766; TF Fork_head. DR Pfam; PF00250; Fork_head; 1. DR PRINTS; PR00053; FORKHEAD. DR ProdOm; PD000425; TF Fork_head; 1. DR SMART; SM00339; FH; 1. DR PROSTE; PS000658; FORK_HEAD-2; 1. DR PROSTE; PS50039; FORK_HEAD-3; 1. DR SEQUENCE 653 AA; 6944 MW; 31B6A6785F4F7E59 CRC64;	
Xiphophorus maculatus (Southern platyfish).	OC Eukaryota; Metazoa; Chordata; Teleostei; Buteleostomi; Neoteleostei; Acanthomorpha; Acanthopterygii; Percomorpha; Atherinomorpha; Cyprinodontiformes; Poeciliidae; Xiphophorus.	Query Match 77.0%; Score 427.5; DB 11; Length 653; Best Local Similarity 73.5%; Pred. No. 8.5e-40; Matches 75; Conservative 18; Mismatches 8; Indels 1; Gaps 1;
[1]	SEQUENCE FROM N.A. STRAIN=Up163_B85_S(A); NCBITaxID=80833; RC Rudd M.D.; Kazanis S.; Butler A.P.; RT From a Xiphophorus Melanoma Model. " ; DR GO; GO:0005634; C:nucleus; - . DR GO; GO:0003700; F:transcription factor activity; IEA. DR InterPro; IPR001766; TF Fork head.	Qy 1 KKTTRRNAGNMYSAYELITTAIMASPEKRILTAQYEMVQNYFRDKGDNSAGWK 60 Db 149 KSSSSRNAGNLSTADLTKAESSAEKRUTLSQIYDMVNTSIPFKDKGDNSAGWK 208
SEQUENCE FROM N.A. STRAIN=Up163_B85_S(A); NCBITaxID=80833; RC Rudd M.D.; Kazanis S.; Butler A.P.; RT From a Xiphophorus Melanoma Model. " ; DR GO; GO:0005634; C:nucleus; - . DR GO; GO:0003700; F:transcription factor activity; IEA. DR InterPro; IPR001766; TF Fork head.	Qy 61 NSIRHNLSHSRMFLQNEGAKGKSSWWVNWPD-AKGGMNPRR 101 Db 209 NSIRHNLSHSRFIRVQNEGIGKSSWWMINPEGGKSKAPRR 250	
RESULT 9 Q8MK69 ID Q8MK69 PRELIMINARY; AC Q8MK69; Q8MK69; PRELIMINARY; AC Q8MK69; DT 01-OCT-2002 ("TREMBlrel. 22, Created)		

Best Local Similarity	73.5%;	Pred.	No. 1.1e-39;
Matches	75;	Conservative	18; Mismatches
			8; Indels 1; Gaps 1
Qy	1 KKTTRNANGNNSYAELITTAIMASPERKLTIAQYEMWVQNPYFRDKGDSNSAGWK 60		
Db	148 KTSSSRNANGNNSYAELITTAIMASPERKLTLSQIYEMWVQNPYFRDKGDSNSAGWK 207		
Qy	61 NSIRHNLSLHSFMRITQNEGAGKSWSWVINPD-AKPGMNPR 101		
Db	208 NSIRHNLSLHSFIRVQNETGCKSSWWMLNPEGGSCKSPR 249		
RESULT 11			
Q9WVH5	PRELIMINARY;	PRT;	652 AA.
ID Q9WVH5;			
AC Q9WVH5;			
DT 01-NOV-1999 (TREMBLrel. 12, Created)			
RA 01-JUN-2003 (TREMBLrel. 12, Last sequence update)			
RT Forkhead protein FKHL.			
DE FOXO1 OR FKHL.			
PRINTER; IPR001766; TF_Fork_head.			
PRINTS; PR00053; FORKHEAD.			
PRODOM; PD000425; TF_Fork_head; 1.			
SMART; SM00339; PH; 1.			
PROSITE; PS00658; FORK_HEAD_2; 1.			
PROSITE; PS50039; FORK_HEAD_3; 1.			
SEQUENCE FROM N.A.			
zhu Q., Cunningham M.A., Hammond J.M., RT FKHL Expression in Porcine Granulosa Cells.;"			
Submitted (APR-2002) to the EMBL/GenBank/DBJ databases.			
RL EMBL; AY094061; BAM19156.1; -.			
DR CO; GO:005634; C:nucleus; IEA.			
DR GO; GO:003700; P:transcription factor activity; IEA.			
DR GO; GO:006355; P:regulation of transcription, DNA-dependent; IEA.			
DR GO; GO:003700; P:transcription factor activity; IEA.			
DR InterPro; IPR001766; TF_Fork_head.			
DR Pfam; PF00250; Fork_head; 1.			
DR PRINTS; PR00053; FORKHEAD.			
DR PRODOM; PD000425; TF_Fork_head; 1.			
DR SMART; SM00339; PH; 1.			
DR PROSITE; PS00658; FORK_HEAD_2; 1.			
DR PROSITE; PS50039; FORK_HEAD_3; 1.			
SEQUENCE 662 AA; 69861 MW;	96ABBB41F5E5E9B CRC64;		
Q9JUW4	PRELIMINARY;	PRT;	652 AA.
ID Q9JUW4;			
AC Q9JUW4;			
DT 01-OCT-2000 (TREMBLrel. 15, Created)			
DT 01-OCT-2000 (TREMBLrel. 15, Last sequence update)			
DT 01-JUN-2003 (TREMBLrel. 24, Last annotation update)			
DR FOXO1 OR FKHL.			
DR FOXO1 OR FKHL.			
Mus musculus (Mouse).			
OC Bony fish.			
OC Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.			
NCBI_TAXID=10090;			
SEQUENCE FROM N.A.			
zhu Q., Cunningham M.A., Hammond J.M., RT FKHL Expression in Porcine Granulosa Cells.;"			
Submitted (APR-2002) to the EMBL/GenBank/DBJ databases.			
RL EMBL; AY094061; BAM19156.1; -.			
DR CO; GO:005634; C:nucleus; IEA.			
DR GO; GO:003700; P:transcription factor activity; IEA.			
DR GO; GO:006355; P:regulation of transcription, DNA-dependent; IEA.			
DR GO; GO:003700; P:transcription factor activity; IEA.			
DR InterPro; IPR001766; TF_Fork_head.			
DR PRINTS; PR00053; FORKHEAD.			
DR PRODOM; PD000425; TF_Fork_head; 1.			
DR SMART; SM00339; PH; 1.			
DR PROSITE; PS00658; FORK_HEAD_2; 1.			
DR PROSITE; PS50039; FORK_HEAD_3; 1.			
SEQUENCE 662 AA; 69861 MW;	96ABBB41F5E5E9B CRC64;		
Q9JUW4	PRELIMINARY;	PRT;	652 AA.
ID Q9JUW4;			
AC Q9JUW4;			
DT 01-OCT-2000 (TREMBLrel. 15, Created)			
DT 01-OCT-2000 (TREMBLrel. 15, Last sequence update)			
DT 01-JUN-2003 (TREMBLrel. 24, Last annotation update)			
DR FOXO1 OR FKHL.			
DR FOXO1 OR FKHL.			
Mus musculus (Mouse).			
OC Bony fish.			
OC Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.			
NCBI_TAXID=10090;			
SEQUENCE FROM N.A.			
Leenders H., Benoist C., Mathis D.; "The forkhead FKHL is involved in thymocyte proliferation."			
Submitted (JAN-2000) to the EMBL/GenBank/DBJ databases.			
EMBL; AY25157; ABAB8873.1; -.			
HSSP; Q63245; 2HFH.			
MGD: MGI:1890077; Foxox1.			
DR GO; GO:005634; C:nucleus; IEA.			
DR GO; GO:003700; P:transcription factor activity; IEA.			
DR GO; GO:006355; P:regulation of transcription, DNA-dependent; IEA.			
DR InterPro; IPR001766; TF_Fork_head.			
DR PRINTS; PR00053; FORKHEAD.			
DR PRODOM; PD000425; TF_Fork_head; 1.			
DR SMART; SM00339; PH; 1.			
DR PROSITE; PS00658; FORK_HEAD_2; 1.			
DR PROSITE; PS50039; FORK_HEAD_3; 1.			
SEQUENCE 652 AA; 69518 MW;	3EFF58636EA85205F CRC64;		
Q9WT78	PRELIMINARY;	PRT;	651 AA.
ID Q9WT78;			
AC Q9WT78;			
DT 01-NOV-1999 (TREMBLrel. 12, Created)			
DT 01-NOV-1999 (TREMBLrel. 12, Last sequence update)			
DR FOXO1 OR FKHL.			
DR FOXO1 OR FKHL.			
Brachydanio rerio (Zebrafish) (Danio rerio).			
OC Actinopterygii; Neotropical; Teleostei; Osteichthyes; Osteopeltiformes.			

OC	Cyprinidae; <i>Danio</i> .
OX	[1] NCBI_TaxID=7935;
RP	SEQUENCE FROM N.A.
RX	MEDLINE=21251166; PubMed=11353388;
RA	Biggs W.H. III, Cavenee W.K., Arden K.C.; "Identification and characterization of members of the FKHR (FOX O) subclass of winged helix transcription factors in the mouse.";
RT	Matlin J., Genome 12:416-425 (2001).
RL	EMBL; AF114262; AAU42110.1; -.
DR	HSSP; Q61245; 2HFH.
DR	TRANSFAC; TO4724; -.
DR	ZFIN; ZDB-GENE-990708-6; foxo5.
GO	GO:0005634; C:nucleus; IEA.
DR	GO; GO:0003700; F:transcription factor activity; IEA.
GO	GO:0006355; P:regulation of transcription, DNA-dependent; IEA.
DR	InterPro; IPR001766; TF_Fork_head.
DR	Pfam; PF00250; Fork_head_1.
PRINTS	PR00053; FORKHEAD.
DR	ProDom; PD000425; TF_Fork_head; 1.
DR	SMART; SM00339; PH; 1.
DR	PROSITE; PS00658; FORK_HEAD_2; 1.
DR	PROSITE; PS50039; FORK_HEAD_3; 1.
SQ	02AO98BB5029C98 CRC64;
	651 AA; 69861 MW;
	Query Match Score 425.5; DB 13; Length 651;
	Best Local Similarity 72.5%; Pred. No. 1.4e-39; Gaps 1;
	Matches 74; Conservative 19; Mismatches 8; Indels 1; Gaps 1;
QY	1 KKTITTRNANGNMSYAELTTAIMASPEKRLTLAQEVVWVNQVNPYFRDGDSNSAGWK 60
Db	118 RKSRRNANGNLSTADLTKIAESTPDKRQLTLSQIYDWNVSSVPYFRDGDSNSAGWK 177
QY	61 NSIRHNLSSLRSRMTIQNEGAGKSSWWMINPEGGGGKAPRR 101
Db	178 NSIRHNLSSLRSRFTIQNECTGKSSWWMINPEGGGGKAPRR 219
RESULT 13	
O16850	PRELIMINARY; PRT; 530 AA.
ID	Q9ES26 PRELIMINARY; PRT; 333 AA.
AC	Q16850; [1]
DT	01-JAN-1998 (TrEMBLrel. 05, Created)
DT	01-JAN-1998 (TrEMBLrel. 05, Last sequence update)
DT	01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE	Fork head-related transcription factor DAF-16b (Hypothetical protein R13H8.1a).
DE	DAF-16 OR R13H8.1.
GN	Caenorhabditis elegans.
OS	Bukaryota; Metazoa; Chordata; Craniata; Vertebrata; Rhabditoidea; Rhabditidae; Peleorinae; Caenorhabditis.
OC	NCBI_TaxID=6239;
OX	[1] SEQUENCE FROM N.A.
RP	SEQUENCE FROM N.A.
RC	STRAIN=N2;
RX	MEDLINE=98013175; PubMed=9351326;
RA	Ogg S., Paradis S., Gottlieb S., Patterson G.I., Lee L., Tissenbaum H.A., Ruvkun G.; "The Fork head domain transmits insulin-like metabolic and longevity signals in <i>C. elegans</i> ."; Nature 389:994-999 (1997).
RT	[2] SEQUENCE FROM N.A.
RP	SEQUENCE FROM N.A.
RC	MEDLINE=99069613; PubMed=9851916;
RA	None; "Genome sequence of the nematode <i>C. elegans</i> : a platform for investigating biology. The <i>C. elegans</i> Sequencing Consortium."; Science 282:2012-2018 (1998).
RT	[3] SEQUENCE FROM N.A.
RN	STRAIN=Bristol N2;
RC	Jones K., Hinds, k, Sutteeer C., Cofman M.;

RT "The sequence of *C. elegans* cosmid R13H8.";
Submitted (JAN-1998) to the EMBL/GenBank/DBJ databases.

RN [4] SEQUENCE FROM N.A.

RP STRAIN=Bristol N2;

RC Waterston R.;

RA "Direct Submission.";

RT "Direct Submission." to the EMBL/GenBank/DBJ databases.

RL Submitted (AUG-2001) to the EMBL/GenBank/DBJ databases.

DR EMBL; AF020344; AA84392.1; -.

DR PIR; T42234; T42234.

DR HSSP; Q61245; 2HFH.

DR TRANSFAC; T03400; -.

DR Wormpep; R13H8.1a; CE2B771.

DR GO; GO:0005634; C:nucleus; IEA.

DR GO; GO:0003700; F:transcription factor activity; IEA.

DR GO; GO:0006355; P:regulation of transcription, DNA-dependent; IEA.

DR InterPro; IPR001766; TF_Fork_head.

DR Pfam; PF00250; Fork_head; 1.

DR PRINTS; PR00053; FORKHEAD.

DR ProDom; PD000425; TF_Fork_head; 1.

DR SMART; SM00339; PH; 1.

DR PROSITE; PS00658; FORK_HEAD_2; 1.

DR PROSITE; PS50039; FORK_HEAD_3; 1.

DR SEQUENCE 530 AA; 57398 MW; 8665FBDB428039D6 CRC64;

Query Match Score 67.1%; DB 5; Length 530;

Best Local Similarity 61.1%; Pred. No. 1.2e-33;

Matches 66; Conservative 15; Mismatches 22; Indels 5; Gaps 1;

QY 1 KKTTTTRNANGNMSYAELTTAIMASPEKRLTLAQEVVWVNQVNPYFRDGDSNS 55

Db 150 KKPDDOLAQKPNPGEESYSDILAKAKLAPAGDGLKLNEYQWPSDNIPYFGERSSPBE 209

QY 56 SAGWKNSIRHNLSSLRSRMTIQNEGAGKSSWWMINPEGGGGKAPRR 103

Db 210 AAGWKNSIRHNLSSLRSRMTIQNEGAGKSSWWMINPEGGGGKAPRR 257

RESULT 14

Q9ES26 PRELIMINARY; PRT; 333 AA.

ID Q9ES26 PRELIMINARY; PRT; 333 AA.

AC Q9ES26; [1]

DT 01-MAR-2001 (TrEMBLrel. 16, Created)

DT 01-MAR-2001 (TrEMBLrel. 16, Last sequence update)

DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)

DE Transcription factor FKH (Fragment).

OS Rattus norvegicus (Rat).

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Rhabditoidea;

OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.

OC NCBI_TaxID=10116;

DR HSSP; Q63245; 2HFH.

DR GO; GO:0005634; C:nucleus; IEA.

DR MEDLINE=20549654; PubMed=10960473;

DR Schmolli D., Walker K.S., Alessi D.R., Grampier R., Burchell A., Guo S., Walther R., Utterman T.G.;

RA "Regulation of Glucose-6-phosphatase gene expression by protein kinase Balpha and the forkhead transcription factor FKH. Evidence for insulin response unit-dependent and -independent effects of insulin on promoter activity.";

RT J. Biol. Chem. 275:36324-36333 (2000).

DR EMBL; AF247812; AAQ03779.1; -.

DR PROSITE; PS00653; FORKHEAD.

DR GO; GO:0003700; F:transcription factor activity; IEA.

DR GO; GO:0006355; P:regulation of transcription, DNA-dependent; IEA.

DR InterPro; IPR001766; TF_Fork_head.

DR Pfam; PF00250; Fork_head; 1.

DR PRINTS; PR00053; FORKHEAD.

DR PRODom; PD000425; TF_Fork_head; 1.

DR SMART; SM00339; PH; 1.

DR PROSITE; PS00658; FORK_HEAD_2; 1.

DR PROSITE; PS50039; FORK_HEAD_3; 1.

PT NON_TER 1 1
 PT NON_TER 333 333 MW: D86C918F0539361D CRC64;
 SQ SEQUENCE 333 AA; 35994 MW: D86C918F0539361D CRC64;

Query Match Score 339.5; DB 11; Length 333;
 Best Local Similarity 75.6%; Pred. No. 3.6e-30;
 Matches 59; Conservative 13; Mismatches 5; Indels 1; Gaps 1;

QY 25 ASPEKRITLAQYQEMMNYQVYFRDKGDSNSAGKNSIRANISLISRFMQLNEAGKS 84
 :|:|||:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:
 Db 1 SAAEKRTLSQYEMMYSVYFKDKGDSNSAGKNSIRANISLISRFMQLNEAGKS 60
 :|:|||:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:
 QY 85 SWWVIND-AKGGMNPPR 101
 :|:||:|:|:|:|:|:|:
 Db 61 SWWMLNPEGKGSKSPRR 78

RESULT 15
 Q9PVK6 PRELIMINARY; PRT; 204 AA.
 ID Q9PVK6
 AC Q9PVK6;
 DT 01-MAY-2000 (T-EMBLrel. 13, Created)
 DT 01-JUN-2000 (T-EMBLrel. 13, Last sequence update)
 DT 01-JUN-2003 (T-EMBLrel. 24, Last annotation update)
 DE Putative fork head domain protein.
 OS Eukaryota; Metazoa; Chordata; (Danio rerio).
 OC Actinopterygii; Neopterygii; Vertebrata; Craniata; Euteleostomi;
 OC Cyprinidae; Danio; Danio.
 OX NCBI_TaxID:7955;
 RN [1]

RP SEQUENCE FROM N.A.
 RC TISSUE=Kidney;
 RA Icard-Liepkalns C, Haire R.N., Strong S.J., Litman G.W.;
 RL Submitted (MAY-1998) to the ENSEMBL/GenBank/DBJ databases.
 DR EMBL: AF064828; AAD53002.1; -.
 DR HSSP: Q63245; 2HFF.
 DR GO: 0005634; C:nucleus; IEA.
 DR GO: 0003700; F:transcription factor activity; IEA.
 DR GO: 0006355; P:regulation of transcription, DNA-dependent; IEA.
 DR InterPro: IPR001766; TF_Fork_head.
 DR Pfam: PF00050; Fork_head_1.
 DR PRINTS: PR000053; FORKHEAD.
 DR PRODOM: PDO0445; TF_Fork_head_1.
 DR SMART: SM00339; FH_1.
 DR PROSITE: PS50019; FORK_HEAD_1.
 DR PROSITE: PS50019; FORK_HEAD_3; 1.
 SQ SEQUENCE 204 AA; 22222 MW; BA78A856937B8CD9 CRC64;

Query Match Score 286; DB 13; Length 204;
 Best Local Similarity 72.1%; Pred. No. 2.3e-24;
 Matches 49; Conservative 15; Mismatches 4; Indels 0; Gaps 0;
 QY 1 KKTTPRNAGNMYSAAELITTAIMASPEKELTIAQYEMMNYQVYFRDKGDSNSAGWK 60
 :|:||:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:
 Db 118 RKSSSERNAQCNLSTADLITKAESTPDKRLTLSQTYDNMVSVPYFKDKDTNSAGWK 177
 :|:||:|:|:|:|:|:|:|:|:
 QY 61 NSIRNLIS 68
 :|:||:|:
 Db 178 YSIRNLIS 185

Search completed: July 23, 2004, 10:25:28
 Job time : 41 secs

This Page Blank (uspto)

Result No.	Score	Query Match	Length	DB ID	Description
1	555	100.0	103	AAB06073	Aab06073 Caenorhabditis elegans
2	555	100.0	3	AAB06075	Aab06075 Caenorhabditis elegans
3	555	100.0	109	AAB06075	Aab06075 Caenorhabditis elegans
4	555	100.0	510	AAB06070	Aab06070 Caenorhabditis elegans
5	517	93.2	97	AAB74317	Aab74317 C. elegans
6	440.5	79.4	673	AAV96449	Aay96449 Forkhead
7	427.5	77.0	655	AAV96448	Aay96448 Forkhead
8	427.5	77.0	655	AAB06076	Aab06076 Human homolog
9	427.5	77.0	656	7 ADD47623	Add47623 Human homolog
10	406.5	77.0	656	7 ADD48792	Add48792 Human homolog
11	422.5	76.1	593	6 ABO14656	Abo14656 Novel human protein
12	410.5	74.0	358	4 AAU00214	Aau00214 Fusion protein
13	410.5	74.0	505	4 AAM39277	Aam39277 Human polypeptide
14	410.5	74.0	513	4 AAM41063	Aam41063 Human polypeptide
15	406.5	73.2	501	3 AAV96447	Aay96447 Forkhead
16	406.5	73.2	501	5 ABG95042	Abg95042 Human transmembrane protein
17	406.5	73.2	501	5 ABG95043	Abg95043 Human transmembrane protein
18	406.5	73.2	505	6 ABP72186	Abp72186 Human transmembrane protein
19	378.5	68.2	509	3 AAB06071	Aab06071 Caenorhabditis elegans
20	378	68.1	98	3 AAB06077	Aab06077 Caenorhabditis elegans
21	367.5	66.2	99	2 AAY55739	Aay55739 EKFR forkhead
22	265	42.7	448	4 ABP59659	Abp59659 Drosophila melanogaster protein
23	222.5	40.1	588	6 AAE37024	Aae37024 Human nucleic acid
24	195	35.1	740	4 ABB66642	Abb66642 Drosophila melanogaster protein
25	195	35.1	740	4 ABB66641	Abb66641 Drosophila melanogaster protein

CC insulin signalling pathway. Conserved DAF motifs can be used to design
 CC probes to identify mammalian DAF homologues and thus to identify
 CC individuals with a predisposition toward the development of glucose
 CC intolerance conditions, such as obesity and diabetes

XX Sequence 103 AA;

Query Match	100.0%	Score 555; DB 3; Length 103;
Best Local Similarity	100.0%	Pred. No. 7.7e-60;
Matches	103;	Conservative 0; Mismatches 0;
Indels	0;	Gaps 0;

QQ 1 KKTTRRNAGNMSYAEPLITTAIASPEKRLLAQVYNNVQNPYFRDKGDSNSAGWK 60

Db 1 KKTTRRNAGNMSYAEPLITTAIASPEKRLLAQVYNNVQNPYFRDKGDSNSAGWK 60

Qy 61 NSIRHNLSLHSRMRQNEGAKSSWWVINDAKPGMNPRTTR 103

Db 61 NSIRHNLSLHSRMRQNEGAKSSWWVINDAKPGMNPRTTR 103

RESULT 2
AAB06065 standard; protein; 106 AA.

XX AAB06065;

AC AAB06065;

XX DE Caenorhabditis elegans DAF-16 forkhead DNA binding domain.

DT 11-JAN-2001 (first entry)

XX Caenorhabditis elegans; daf-16; daf-18; insulin signalling pathway;

XX daf-2; age-1; insulin receptor; PI 3-kinase; PKB kinase; AKT kinase;

XX PTEN lipid phosphatase; antidiabetic; anorectic; obesity; diabetes;

XX KW impaired glucose tolerance; transgenic animal; conserved motif.

XX OS Caenorhabditis elegans.

XX PN WO200033068-A1.

XX PD 08-JUN-2000.

XX PR 02-DEC-1999; 99WO-US028529.

XX PR 03-DEC-1998; 98US-00205658.

XX PA (GEHO) GEN HOSPITAL CORP.

XX PI Ruvkun G, Ogg S;

XX DR WPI; 2000-423022/36.

XX Disclosure; Page 134; 402pp; English.

XX PS Diagnosing and treating obesity and impaired glucose tolerance using

XX modulators of daf-18 expression and/or activity.

XX Disclosure; Page 134; 402pp; English.

XX DR WPI; 2000-423022/36.

XX Diagnosing and treating obesity and impaired glucose tolerance using

XX modulators of daf-18 expression and/or activity.

XX Disclosure; Page 134; 402pp; English.

QY	1	KKTTTRNAGNMSYAELITTAIMASPERKLTLAQVYEMVQNYPYFRDKGDSNSSAGWK	XX	Sequence 510 AA;
Db	7	KKTTTRNAGNMSYAELITTAIMASPERKLTLAQVYEMVQNYPYFRDKGDSNSSAGWK	60	Query Match 100.0%; Score 555; DB 3; Length 510;
QY	61	NSIRHNLSLHSRMRNQNGAGSSWWINPDAKPGMNRTRR 103	XX	Best Local Similarity 100.0%; Pred. No. 6.2e-59;
Db	67	NSIRHNLSLHSRMRNQNGAGSSWWINPDAKPGMNRTRR 109	60	Matches 103; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
RESUME 4				
AAB06070	ID	AAB06070 standard; protein; 510 AA.	Db	1 KKTTRNAGNMSYAELITTAIMASPERKLTLAQVYEMVQNYPYFRDKGDSNSSAGWK 60
AC			Db	135 KKTTTRNAGNMSYAELITTAIMASPERKLTLAQVYEMVQNYPYFRDKGDSNSSAGWK 194
AAE06070	XX		QY	61 NSIRHNLSLHSRMRNQNEGAKSSWWINPDAKPGMNRTRR 103
AC	XX		Db	195 NSIRHNLSLHSRMRNQNEGAKSSWWINPDAKPGMNRTRR 237
XX	DT	11-JAN-2001 (first entry)	RESULT 5	
XX	DE	Caenorhabditis elegans DAF-16 isoform #1.	ID	AAB74317 standard; peptide; 97 AA.
XX	KW	Caenorhabditis elegans; daf-16; daf-2; age-1; daf-18;	XX	AAB74317;
XX	KW	insulin signalling pathway; insulin receptor; PI 3-kinase; PKB kinase;	XX	27-JUN-2001 (first entry)
XX	KW	AKT kinase; PTEN lipid phosphatase; antidiabetic; anorectic; obesity;	XX	
XX	KW	diabetes; impaired glucose tolerance; transgenic animal.	DE	C. elegans DAF-16 winged helix domain.
OS			XX	DAF-16; winged helix domain; homolog; ageing; ischemia; viral;
OS			XX	KW cardiac hypertrophy; fever; inflammation; metabolic disease; bacterial; infection; cancer; tumour; stress.
OS			XX	Caenorhabditis elegans.
XX	FH	Key Location/Qualifiers	XX	DAF-16; winged helix domain; homolog; ageing; ischemia; viral;
XX	FT	Misc-difference 119	XX	KW cardiac hypertrophy; fever; inflammation; metabolic disease; bacterial; infection; cancer; tumour; stress.
XX	FT	/note= "encoded by CTR"	XX	XX
XX	FT	Misc-difference 120	XX	OS
XX	FT	/note= "encoded by CGG"	XX	XX
XX	FT	Misc-difference 121	XX	XX
XX	FT	/note= "encoded by AGT"	XX	XX
XX	FT	295	PN	WO200118549-A1.
XX	FT	/note= "encoded by GA"	PR	99US-0152825P.
XX	XX		PD	15-MAR-2001.
XX	XX		PF	07-SEP-2000; 20000WO-US024487.
XX	PN	WO200033068-A1.	XX	XX
XX	PD	08-JUN-2000.	PR	07-SEP-1999;
XX	XX		XX	(NEUR-) NEUROGENETICS INC.
XX	PF	99WO-US028529.	PA	
XX	PR	02-DEC-1999;	XX	PI Henderson ST, Johnson TE;
XX	PR	03-DEC-1998;	XX	XX
XX	PN	98US-00205658.	DR	WPI; 2001-226755/23.
XX	PA	(GEHO) GEN HOSPITAL CORP.	XX	XX
XX	PD	08-JUN-2000.	PT	Identifying agents that can increase the activity of DAF-16 or their
XX	XX		PT	mammalian homologs, useful for inhibiting tumor growth and for protecting
XX	PF	99WO-US028529.	PT	the cells of an organism from environmental stress, e.g. heat and
XX	PR	02-DEC-1999;	PT	cytotoxic drugs.
XX	PR	03-DEC-1998;	PS	Disclosure: Page 10; 59pp; English.
XX	XX	98US-00205658.	CC	The present invention relates to methods of identifying agents that can
XX	PA	(GEHO) GEN HOSPITAL CORP.	CC	increase the activity of C.elegans DAF-16 or their mammalian homologs.
XX	PI	Ruvkun G, Ogg S;	CC	Compounds identified by the methods are useful in any disease state where
XX	PS	02-DEC-1999;	CC	tissue damage results or in ageing, for inhibiting certain tumour growth
XX	DR	99WO-US028529.	CC	and for protecting the cells of an organism from damage associated with
DR	DR	03-DEC-1998;	CC	diseases, e.g. ischemia, cardiac hypertrophy, fever, inflammation,
N-PSDE	N-PSDE	98US-00205658.	CC	metabolic diseases, viral and bacterial infection, cell and tissue
XX	PT	Diagnosing and treating obesity and impaired glucose tolerance using	CC	trauma, and cancer. The compounds are also useful for protecting the
XX	PT	modulators of daf-18 expression and/or activity.	CC	cells of an organism from environmental stress. The present sequence is
XX	PS	Disclosure: Fig 14A; 402pp; English.	CC	the C. elegans DAF-16 winged helix domain. This peptide contains a
XX	PS	The present cDNA sequence from Caenorhabditis elegans corresponds to a	CC	putative DNA binding region.
XX	PS	CC differentially spliced daf-16 transcript which encodes an isoform of DAF-16. DAF-16 interacts with the C. elegans insulin signalling pathway. A	CC	Sequence 97 AA;
CC	CC	number of C. elegans genes have been identified as homologues of genes in	CC	Query Match 93.2%; Score 517; DB 4; Length 97;
CC	CC	the mammalian insulin signalling pathway. The C. elegans age-1 gene	CC	Best Local Similarity 100.0%; Pred. No. 3.1e-55;
CC	CC	encodes a homologue of the mammalian PI 3-kinase whilst daf-2 encodes a	CC	Matches 96; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
CC	CC	homologue of the mammalian insulin receptor. The C. elegans PI 3-kinase	CC	
CC	CC	homologues act downstream of insulin signalling. Other daf genes have	CC	
CC	CC	also been implicated in the C. elegans insulin signalling pathway. The C.	CC	
CC	CC	elegans PTEN lipid phosphatase homologue, DAF-18, has been found to act	CC	
CC	CC	upstream of AKT in the pathway. This discovery has enabled mammalian PTEN	CC	
CC	CC	action to be mapped to the insulin signalling pathway. Compounds that	CC	
CC	CC	inhibit the expression and/or activity of polypeptides encoded by these	CC	
CC	CC	genes may be administered to patients to treat or prevent disorders such	CC	
CC	CC	as obesity and impaired glucose tolerance	CC	

Db	1	KKTTTRNAWGMSYAELETTAIMASPEKRITLAQVYEMMVNQVYFRDKGDSNSSAGWK	60	RESULT 7
Qy	61	NSIRHNLSLHSRMRIONEGAKSSWWVINPDAKG 96		AAV96448 Standard; protein: 655 AA.
Db	61	NSIRHNLSLHSRMRIONEGAKSSWWVINPDAKG 96		ID AAV96448
			XX	XX
			AC	AAV96448;
			XX	XX
			DT	12-SEP-2000 (first entry)
			DE	Forkhead transcription factor FKHR.
			XX	XX
			XX	Protein kinase B; PKB; substrate; phosphorylation; forkhead; Afx; FKHR.
			KW	KW
			KW	transcription factor; insulin; suppressor.
			XX	XX
			OS	Homo sapiens.
			XX	XX
			PN	WO200031291-A1.
			XX	
			PD	02-JUN-2000.
			XX	02-JUN-2000.
			PF	16-NOV-1999;
			XX	99WO-SE002095.
			XX	
			PR	19-NOV-1998;
			XX	98SE-00003958.
			PA	(PHAA) PHARMACIA & UPJOHN AB.
			XX	
			PI	James S, Dahlman-Wright K, Lake S, Butcher S, Climent I;
			XX	
			DR	WPI; 2000-400095/34.
			XX	
			PT	Method for screening substances which modulate Protein kinase B activity
			PT	uses substrate peptides.
			XX	XX
			PS	Disclosure; Page 16-19; 41pp; English.
			XX	XX
			CC	This is the forkhead transcription factor, FKHR, which is related to the
			CC	Caenorhabditis elegans nematode worm forkhead protein, daf-16. FKHR is
			CC	phosphorylated by protein kinase B (PKB). FKHR is related to Afx (AAV96447), which is also phosphorylated by PKB. This phosphorylation, in
			CC	CC
			CC	enhanced when cells are stimulated with insulin or IGF-1. PKB
			CC	therefore demonstrates a novel regulation of transcription. Afx-
			CC	CC
			CC	mediated stimulation of IRE driven reporter genes is suppressed by
			CC	CC
			CC	overexpression of PKB and this is further suppressed by insulin. This
			CC	CC
			CC	mechanism can be used to identify compounds for treatment of metabolic
			CC	CC
			CC	disease. Substrate peptides (AAV96441-46) derived from Afx, and forkhead
			CC	CC
			CC	proteins FKHR and FKHL1 (AAV96449), can be used to discriminate between
			CC	CC
			CC	the effects of compounds which mediate insulin action through
			CC	CC
			CC	transcription via forkhead transcription factor family from those which
			CC	CC
			CC	modulate activity of enzymes involved in metabolism by phosphorylation.
			CC	CC
			CC	The claimed method comprises screening for substances which are
			CC	CC
			CC	activators, inhibitors or binders of PKB by the use of 2 sequences chosen
			CC	CC
			CC	from AAY96441 and AAY96442, AAY96643 and AAY96644 and AAY96645 and
			CC	CC
			AAY96646	AAY96646
			XX	Sequence 655 AA;
			SQ	Query Match 75; Conservative 18; Mismatches 8; Indels 1; Gaps 1;
				Best Local Similarity 73.5%; Pred. No. 3.7e-43;
				Matches 75; Score 427.5; DB 3; Length 655;
Qy	1	KKTTTRNAWGMSYAELETTAIMASPEKRITLAQVYEMMVNQVYFRDKGDSNSSAGWK	60	RESULT 8
Db	151	NSIRHNLSLHSRMRIONEGAKSSWWVINPDAKG 96		AB06076
			ID	AB06076 standard; protein: 655 AA.
			XX	
			AC	AB06076;
Qy	1	KKTTTRNAWGMSYAELETTAIMASPEKRITLAQVYEMMVNQVYFRDKGDSNSSAGWK	60	RESULT 8
Db	148	RKSSSRRAWGMSYADLITRAESTSPDKRLTLSQTEWMVRCPYFRDKGKSCKSPRR	249	AB06076
			ID	AB06076 standard; protein: 655 AA.
			XX	
			AC	AB06076;

XX	11-JAN-2001	(first entry)	KW	Human; pain; neuronal tissue; gene therapy;
XX	Human homologue of Caenorhabditis elegans DAF-16.		KW	spinal segmental nerve injury; chronic constriction injury; CCI;
DE			KW	spared nerve injury; SNI; Chung.
XX	Human; Caenorhabditis elegans; daf-16; daf-18;		XX	
KW	insulin signalling pathway; daf-2; age-1; insulin receptor; PI 3-kinase;		OS	Homo sapiens.
KW	PKB kinase; AKT kinase; PTEN lipid phosphatase; antidiabetic; anorectic;		PN	WO2003016475-A2.
KW	obesity; diabetes; transgenic animal.		XX	
XX			PD	27-FEB-2003.
OS	Homo sapiens.		XX	
XX	W0200303068-A1.		PF	14-AUG-2002; 2002WO-US25765.
PN			XX	
XX	08-JUN-2000.		PR	14-AUG-2001; 2001US-0312147P.
PD			PR	01-NOV-2001; 2001US-0346382P.
XX	02-DEC-1.999;	99WO-US28529.	PR	26-NOV-2001; 2001US-0333347P.
PF			XX	
XX	03-DEC-1.998;	98US-00205658.	PA	(GEHO) GEN HOSPITAL CORP.
PR			PA	(FARB) BAYER AG.
XX	(GEHO) GEN HOSPITAL CORP.		XX	
PA			PI	Woolf C, D'urso D, Beffort K, Costigan M;
XX			XX	
PT	Ruykun G, Ogg S;		DR	WPI; 2003-268312/26.
XX			DR	GENBANK; U02310.
DR	2000-423022/36.		XX	
PT	Diagnosing and treating obesity and impaired glucose tolerance using		PT	New composition comprising two or more isolated polypeptides, useful for preparing a medicament for treating pain in an animal.
XX	modulators of daf-18 expression and/or activity.		XX	
PS	Disclosure; Page 312; 402pp; English.		PS	Claim 1; Page; 1017PP; English.
XX			XX	
CC	The present sequence is the human homologue of the DAF-16 protein from Caenorhabditis elegans. DAF-16 is the major transcriptional output of the C. elegans insulin signalling pathway. A number of C. elegans genes have been identified as homologues of genes in the mammalian insulin signalling pathway. The C. elegans age-1 gene encodes a homologue of the mammalian PI 3-kinase whilst daf-2 encodes a homologue of the mammalian insulin receptor. The C. elegans PKB kinase and AKT kinase act downstream of daf-2 and age-1, just as their mammalian homologues act downstream of insulin signalling. Other daf genes have also been implicated in the C. elegans insulin signalling pathway. The C. elegans PTEN lipid phosphatase homologue, DAF-18, has been found to act upstream of AKT in the pathway. This discovery has enabled mammalian PTEN action to be mapped to the insulin signalling pathway. Conserved DAF motifs can be used to design probes to identify mammalian DAF homologues and thus to identify individuals with a predisposition toward the development of glucose intolerance conditions, such as obesity and diabetes		CC	The invention discloses a composition comprising two or more isolated rat or human polynucleotides or a polynucleotide which represents a fragment, derivative or allelic variation of the nucleic acid sequence. Also claimed are a vector comprising the novel polynucleotide, a host cell comprising the vector, a method for identifying a polynucleotide sequence which is differentially regulated in an animal subjected to pain and a kit to perform the method, an array, a method for identifying an agent that increases or decreases the expression of the polynucleotide sequence that is differentially expressed in neuronal tissue of a first animal subjected to pain, a method for identifying a compound which regulates the expression of a polynucleotide sequence which is differentially expressed in an animal subjected to pain, method for identifying a compound that regulates the activity of one or more of the polynucleotides, a method for producing a pharmaceutical composition, a method for identifying a compound or small molecule that regulates the activity in an animal of one or more of the polypeptides given in the specification, a method for preparing a medicament useful in treating pain and a pharmaceutical composition comprising the one or more polypeptides or their antibodies. The polynucleotide or the compound that modulates its activity is useful for preparing a medicament for treating pain (e.g. spinal segmental nerve injury (SNI), chronic constriction injury (CCI) and spared nerve injury (SNI)) in an animal (e.g. gene therapy). The sequence presented is a human protein (shown in Table 2 of the specification) which is differentially expressed during pain. Note: The sequence data for this patent did not form part of the printed specification, but was obtained in electronic form directly from WIPO at ftp.wipo.int/pub/published_pct_sequences.
SQ	Sequence 655 AA;		SQ	Sequence 656 AA;
Query Match	77.0%	Score 427.5;	DB 3;	Length 655;
Best Local Similarity	73.5%;	Pred. No. 3.7e-43;		
Matches	75;	Conservative 18; Mismatches 8;		
Qy	1	KKTTRNAGNMSAYELTTAIMASPEKRLLTAQYEMMYCNVPYFRDKGDSNSAGWK 60	Qy	1 KKTTRNAGNMSAYELTTAIMASPEKRLLTAQYEMMYCNVPYFRDKGDSNSAGWK 60
Db	151	KSSSSRNAGNLSYADLITKAESSERLLTSQIYEMWYKSPVFKRGCDNSAGWK 210	Db	151 KSSSSRNAGNLSYADLITKAESSERLLTSQIYEMWYKSPVFKRGCDNSAGWK 210
Qy	61	NSIRHNLSLHSRMRMIONEAGKSSWWINPD-AKGGMNPRR 101	Qy	61 NSIRHNLSLHSRMRMIONEAGKSSWWINPD-AKGGMNPRR 101
Db	211	NSIRHNLSLHSRMRMIONEAGKSSWWINPD-AKGGMNPRR 252	Db	211 NSIRHNLSLHSRMRMIONEAGKSSWWINPD-AKGGMNPRR 252
RESULT 9				
ADD47623				
ID	ADD47623	standard; protein; 656 AA.		
XX				
AC	ADD47623;			
XX				
DT	29-JAN-2004 (first entry)			
XX				
DE	Human Protein U02310, SEQ ID NO 13319.			
XX				

ADD48792 standard; protein; 656 AA.
 ID ADD48792;
 XX
 AC ADD48792;
 XX
 DT 29-JAN-2004 (first entry)
 XX
 Human Protein U02310, SEQ ID NO 14502.
 DE
 XX
 KW Human; pain; neuronal tissue; gene therapy;
 KW spinal segmental nerve injury; chronic constriction injury; CCI;
 spared nerve injury; SNI; Chung.
 XX
 OS Homo sapiens.
 XX
 PN WO2003016475-A2.
 XX
 PD 27-FEB-2003.
 XX
 PF 14-AUG-2002; 2002WO-US025765.
 XX
 PR 14-AUG-2001; 2001US-0312147P.
 PR 01-NOV-2001; 2001US-0344382P.
 PR 26-NOV-2001; 2001US-0333347P.
 XX
 PA (GEHO) GEN HOSPITAL CORP.
 PA (FARB) BAYER AG.
 XX
 PI Woolf C, D'urso D, Beffort K, Costigan M;
 XX
 DR 2003-268312/26.
 XX
 DR GENBANK; U02310.
 XX
 PT New composition comprising two or more isolated polypeptides, useful for
 preparing a medicament for treating pain in an animal.
 XX
 PS Claim 1; Page: 1017pp; English.
 XX
 The invention discloses a composition comprising two or more isolated rat
 or human polynucleotides or a polynucleotide which represents a fragment,
 derivative or allelic variation of the nucleic acid sequence. Also
 claimed are vector comprising the novel polynucleotide, a host cell
 comprising the vector, a method for identifying a nucleotide sequence
 which is differentially regulated in an animal subjected to pain and a
 kit to perform the method, an array for identifying an agent
 that increases or decreases the expression of the polynucleotide sequence
 that is differentially expressed in neuronal tissue of a first animal
 subjected to pain, a method for identifying a compound which regulates
 the expression of a polynucleotide sequence which is differentially
 expressed in an animal subjected to pain, a method for identifying a
 compound that regulates the activity of one or more of the
 polynucleotides, a method for producing a pharmaceutical composition,
 a method for identifying a compound or small molecule that regulates the
 activity in an animal of one or more of the polypeptides given in the
 specification, a method for identifying a compound useful in treating
 pain and a pharmaceutical composition comprising the one or more
 polypeptides or their antibodies. The polynucleotide or the compound that
 modulates its activity is useful for preparing a medicament for treating
 pain (e.g. spinal segmental nerve injury (Chung), chronic constriction
 injury (CCI) and spared nerve injury (SNI) in an animal (e.g. gene
 therapy). The sequence presented is a human protein (shown in Table 2 of
 the specification) which is differentially expressed during pain. Note:
 the sequence data for this patent did not form part of the printed
 specification, but was obtained in electronic form directly from WIPO at
 ftp.wipo.int/pub/published_pct_sequences.
 XX
 Sequence 656 AA;

Query Match 77.0%; Score 427.5; DB 7; Length 656;
 Best Local Similarity 73.5%; Pred. No. 3..7e-43;
 Matches 18; Mismatches 8; Indels 1; Gaps 1;
 XX
 1 KKTTRRNAWGMSYAAELITAIMASPERKRLLAQVEMVQPYFRDKGDSNNSAGWK 60

151 KSSSSRRNAWGNLTYADLITKAETTSAEKRLLTUSQIVMVKSPYFRDKGDSNNSAGWK 21.0
 61 NSTRHNLSLHSRMRIONEGAGKGSWWINPD AKPGMNPR 101
 211 NSTRHNLSLHSKFTRVQNEGTTGSSWWMLNPEGKSGKSPRR 252

Db AB014656 standard; protein; 593 AA.
 Qy ID AB014656
 XX
 Db AC AB014656;
 XX
 DT 25-AUG-2003 (first entry)
 XX
 DE Novel human protein #29.
 XX
 KW Human; NOV; gene therapy; endocrine related disease; diabetes;
 KW metabolic-related disease; obesity; central nervous system disorder;
 KW Alzheimer's disease; Parkinson's disease; epilepsy; multiple sclerosis;
 KW schizophrenia; depression; autoimmune disorder; inflammatory disorder;
 KW psoriasis; allergy; lupus erythematosus; asthma; cancer;
 KW inflammatory bowel disease; rheumatoid arthritis; osteoarthritis;
 KW colon cancer; lung cancer; liver cancer; breast cancer; ovarian cancer;
 KW prostate cancer; brain cancer; melanoma; liver disease; haemophilia;
 KW lung disease; emphysema; obstructive pulmonary disease; haemophilia;
 KW stroke; infection.

XX
 Homo sapiens.
 XX
 OS Homo sapiens.
 XX
 PN WO2003023002-A2.
 XX
 PD 20-MAR-2003.
 XX
 PF 09-SEP-2002; 2002WO-US028539.
 XX
 PR 07-SEP-2001; 2001US-0318120P.
 PR 07-SEP-2001; 2001US-0318130P.
 PR 10-SEP-2001; 2001US-0318430P.
 PR 10-SEP-2001; 2001US-032236P.
 PR 17-SEP-2001; 2001US-0322781P.
 PR 17-SEP-2001; 2001US-0322816P.
 PR 17-SEP-2001; 2001US-0322817P.
 PR 17-SEP-2001; 2001US-0323519P.
 PR 19-SEP-2001; 2001US-0323519P.
 PR 20-SEP-2001; 2001US-0323631P.
 PR 20-SEP-2001; 2001US-0323636P.
 PR 25-SEP-2001; 2001US-0324569P.
 PR 25-SEP-2001; 2001US-0325091P.
 PR 26-SEP-2001; 2001US-0324590P.
 PR 17-APR-2002; 2002US-0373212P.
 PR 06-SEP-2002; 2002US-00236177.

(CURA-) CURAGEN CORP.

XX
 PA Spytak KA, Pattrulanj M, Gorman L, Li L, Anderson DW, Zhong M;
 PI Gerlach VL, Vernet CM, Ellerman K, Bergs C, Rothenberg ME, Guo X;
 PI Shimkets RA, Leach MD, Kekuda R, Ji W, Miller CB;
 PI Rieger DK, Taupier RJ, Shenoy SG, Liu X, Padigaru M, Alsobrook JP;
 PI Lepley DM, Edinger SR, Burgess CE;
 XX
 DR WPI; 2003-113242/30.
 DR N-PSDB; ACD1349.
 XX
 PT New cytoplasmic, nuclear membrane bound or secreted polypeptides (NOVX)
 PT and polynucleotides, useful in gene therapy, e.g. for treating or
 PT preventing obesity, multiple sclerosis, allergy, cancers, hemophilia,
 PT stroke or infections.

XX
 PS Claim 1; Page 145; 586pp; English.

XX
 CC The invention describes a new isolated polypeptide (NOVX). The NOVX
 CC polypeptide, nucleic acid and antibody are useful as therapeutics,

CC particularly in the manufacture of a medicament for treating a syndrome
 CC associated with a human disease, which includes a pathology associated
 CC with NOVX polypeptide. The DNA encoding the protein is useful in gene
 CC therapy for treating the disease or condition. In particular, the NOVX
 CC polypeptide or polynucleotide is useful for treating endocrine/
 CC metabolism-related diseases (e.g. obesity or diabetes), central nervous
 CC system disorders (e.g. Alzheimer's disease, Parkinson's disease,
 CC epilepsy, multiple sclerosis, schizophrenia or depression), autoimmune
 CC and inflammatory disorders (e.g. psoriasis, allergy, lupus erythematosus,
 CC asthma, inflammatory bowel disease, rheumatoid arthritis or
 CC osteoarthritis), cancers (e.g. colon, lung, liver, breast, ovarian,
 CC prostate or brain cancers, or melanoma), liver diseases (e.g. liver
 CC cirrhosis), lung diseases (emphysema or obstructive pulmonary disease),
 CC haemophilia, stroke, or infections (e.g. viral, bacterial or parasitic).
 CC These are also useful in developing powerful assay system for functional
 CC analysis of various human disorders, as well as in diagnostic
 CC applications, and for monitoring the effects of drugs during clinical
 CC trials. This is the amino acid sequence of a novel human NOV protein
 XX Sequence 593 AA;

Query Match 76.1%; Score 422.5; DB 6; Length 593;
 Best Local Similarity 80.0%; Pred. No. 1.3e-42; Indels 1; Gaps 1;

Matches 76; Conservative 10; Mismatches 8; Indels 1; Gaps 1;

Qy 8 NAWGNMSYBLLTAIMASPEKRITLAQVYENNYQNVPYFRDKGDSNSAGWNSIRHNL 67
 Db 75 NAWGNLSYDILITPAIESSPDKRITLSQVYENNYQNVPYFRDKGDSNSAGWNSIRHNL 134

Qy 68 SLHSRFMRVLIQNEGAGKSSMWINPD-AKEGMNPRR 101
 Db 135 SLHSRFMRVLIQNEGAGKSSMWINPDGGKSKGAKPRR 169

RESULT 12
 AAU00214 ID AAU00214 standard, protein; 358 AA.

AC AAU00214;

XX DT 09-MAY-2001 (first entry)

XX DE Fusion protein GST-AfxDBD, amino acid sequence.

XX KW DNA binding domain; Afx; transcription factor; GST-AfxDBD:

XX KW human fork head transcription factor; Afx response element; human;

XX KW diabetes; drug target; insulin; insulin receptor signalling pathway.

OS Homo sapiens.

OS Synthetic.

OS Chimeric.

XX FH Key

FT Misc-difference 28 Location/Qualifiers

FT FT note= "Encoded by TAT"

FT FT Misc-difference 349 /note= "Encoded by AG"

FT FT /note= "Encoded by AG"

PN PN WO200114544-A1.

XX PD 26-JUL-2001.

XX PP 26-DEC-2000; 2000WO-US034263.

XX PR 23-DEC-1999; 99US-00471275.

PR 21-JAN-2000; 2000US-0048725.

PR 25-APR-2000; 2000US-0055317.

PR 20-JUN-2000; 2000US-0059042.

PR 19-JUL-2000; 2000US-0062312.

PR 03-AUG-2000; 2000US-00653450.

PR 14-SEP-2000; 2000US-0066191.

PR 19-OCT-2000; 2000US-00693036.

PR 29-NOV-2000; 2000US-00727344.

XX PA (HYSEQ -) HYSEQ INC.

DR N-PSDB; AAS00244.

XX PT New Afx response element with a nucleotide sequence comprising a DNA
 PT binding site for the human fork head transcription factor Afx, useful in
 PT screening for genes or in bioinformatic analysis of the human genome.

XX PS Example 2; Fig 7; 34pp; English.

XX The sequence represents the amino acid sequence of the GST-AfxDB fusion
 CC protein encoded by expression vector pGEX-AfxDB. Human fork head
 CC transcription factor, Afx, was expressed and the protein used to find a
 CC response element comprising an 8 base pair (bp) nucleotide sequence. The
 CC nucleotide sequence comprises AACATGTT, the Afx response element, and is
 CC useful in bioinformatic analysis e.g. of the human genome. Employing the
 CC Afx response element is also useful for screening for genes that may be
 CC used as diabetes drug targets. This can provide a subset of genes
 CC transcriptionally responsive to insulin and may lead to development of
 CC assays that facilitate the analysis of genes interacting with the insulin
 CC receptor pathway. The genes found in such screening may also be employed
 CC in further screening methods for compounds which modify the insulin
 CC receptor signalling pathway

XX SQ Sequence 358 AA;

XX Query Match 74.0%; Score 410.5; DB 4; Length 358;

XX Best Local Similarity 71.6%; Pred. No. 2e-41;

XX Matches 73; Conservative 15; Mismatches 13; Indels 1; Gaps 1;

Qy 1 KKTTTERRNAWGNMSYAEPLITTAIMASPEKRITLAQVYENNYQNVPYFRDKGDSNSAGWNSIRHNL 60

Db 239 RKGGSRRNAWGNQSYAEPLITSAQALESPAKRITLAQVYENNYQNVPYFRDKGDSNSAGWNL 298

Qy 61 NSTRHNLSLHSREPMRIVLQEAGKSSMWVNPD-AKEGMNPRR 101

Db 299 NSTRHNLSLHSKFIVNEATGKSSWMNPEGGKSKAPRR 340

RESULT 13

AAU39277

ID AAU39277 standard; protein; 505 AA.

XX AC AAU39277;

XX DT 22-OCT-2001 (first entry)

XX DB Human polypeptide SEQ ID NO 2422.

XX KW Human; nootropic; immunosuppressant; cytostatic; gene therapy; cancer;
 KW peripheral nervous system; neuropathy; central nervous system; CNS;
 KW Alzheimer's; Parkinson's disease; Huntington's disease; haemostatic;
 KW amyotrophic lateral sclerosis; Shy-Drager Syndrome; chemokinetic;
 KW chemokinetic; thrombolytic; drug screening; arthritis; inflammation;
 KW leukaemia.

XX OS Homo sapiens.

XX PN WO200153312-A1.

XX PD 26-JUL-2001.

XX PP 26-DEC-2000; 2000WO-US034263.

XX PR 23-DEC-1999; 99US-00471275.

PR 21-JAN-2000; 2000US-0048725.

PR 25-APR-2000; 2000US-0055317.

PR 20-JUN-2000; 2000US-0059042.

PR 19-JUL-2000; 2000US-0062312.

PR 03-AUG-2000; 2000US-00653450.

PR 14-SEP-2000; 2000US-0066191.

PR 19-OCT-2000; 2000US-00693036.

PR 29-NOV-2000; 2000US-00727344.

XX PA (HYSEQ -) HYSEQ INC.

XX Tang YT, Liu C, Asundi V, Chen R, Ma Y, Qian XB, Ren F, Wang D; PR 19-JUL-2000; 2000US-00620312.
 PT PR 03-AUG-2000; 2000US-00653450.
 PI PR 14-SEP-2000; 2000US-00662191.
 PI PR 19-OCT-2000; 2000US-00693036.
 XX PR 29-NOV-2000; 2000US-00727344.
 DR XX
 WPI; 2001-442253/47.
 DR N-PSDB; AAI58433.
 XX PA (HYSBQ INC.
 XX Novel nucleic acids and polypeptides, useful for treating disorders such as central nervous system injuries.
 PT PI Tang YT, Liu C, Asundi V, Chen R, Ma Y, Qian XB, Ren F, Wang D; PI Wang J, Wang Z, Wehrman T, Xu C, Xue AJ, Yang Y, Zhang J, Zhao QA;
 PT PI Zhou P, Goodrich R, Drmanac RT; PI Zhou P, Goodrich R, Drmanac RT;
 XX XX
 PS XX
 Example 4; SEQ ID NO 2422; 10078pp; English.
 XX
 CC The invention relates to human nucleic acids (AAI57798-AAI61369) and the encoded polypeptides (AAI38642-AAI42213) with nootropic, immunosuppressant and cytostatic activity. The polynucleotides are useful in gene therapy. A composition containing a polypeptide or polynucleotide system, such as peripheral nervous injuries, peripheral neuropathy and localised neuropathies and central nervous system diseases, such as Alzheimer's, Parkinson's disease, Huntington's disease, amyotrophic lateral sclerosis, and Shy-Drager Syndrome. Other uses include the utilisation of the activities such as: Immune system suppression, Activin/inhibin activity, chemotactic/chemokinetic activity, haemostatic and thrombolytic activity, cancer diagnosis and therapy, drug screening, assays for receptor activity, arthritis and inflammation, leukaemias and C.N.S disorders. Note: The sequence data for this patient did not form part of the printed specification
 XX Sequence 505 AA;
 Query Match Score 410 5; DB 4; Length 505;
 Best Local Similarity 71.6%; Pred. No. 3.2e-41;
 Matches 73; Conservative 15; Mismatches 13; Indels 1; Gaps 1;
 CC
 Qy 1 KKTTPRANQNMSTAYLITTAIMASPEKRLTQAQYEMWVNPYFRDKGDNSSSAGWK 60
 Db 61 NSIRHNLSLHSRMRITLQNEGAKGKSSWVINED-AKEGMNPR 101
 Qy 92 RKGGSRANQNOSTAELISQASAEPEKRLTQAQYEMWRTVPPFKDGDNSSSAGWK 193
 Db 152 NSIRHNLSLHSKFKIVNEATGKSSWMLNPEGGSGKAPRR 193
 XX
 RESULT 14
 AAM41063
 ID AAM41063 standard; protein; 513 AA.
 XX
 AC AAM41063;
 OS Homo sapiens.
 DT 22-OCT-2001 (first entry)
 XX Human polypeptide SEQ ID NO 5994.
 DE AAY96447 standard; protein; 501 AA.
 XX
 KW Human; nootropic; immunosuppressant; cytostatic; gene therapy; cancer;
 KW peripheral nervous system; neuropathy; central nervous system; CNS;
 KW Alzheimer's disease; Huntington's disease; Shy-Drager Syndrome; haemostatic;
 KW amyotrophic lateral sclerosis; drug screening; arthritis; inflammation;
 KW chemoattractant; thrombolytic; chemokinetic; leukaemia.
 XX
 PN WO200153312-A1.
 AX PD 26-JUL-2001.
 XX
 PF 26-DEC-2000; 2000WO-US034263.
 XX
 PR 23-DEC-1999; 900US-00471725.
 PR 21-JAN-2000; 2000US-00488725.
 PR 21-APR-2000; 2000US-00552317.
 PR 20-JUN-2000; 2000US-00598042.
 PR 19-NOV-1998; 98SE-0003958.

XX Novel nucleic acids and polypeptides, useful for treating disorders such as central nervous system injuries.
 PT PT
 XX XX
 PS PS Example 2; SEQ ID NO 5994; 10078pp; English.
 XX XX
 CC The invention relates to human nucleic acids (AAI57798-AAI61369) and the encoded polypeptides (AAI38642-AAI42213) with nootropic, immunosuppressant and cytostatic activity. The polypeptides are useful in gene therapy. A composition containing a polypeptide or polynucleotide of the invention may be used to treat diseases of the peripheral nervous system, such as peripheral nervous injuries, peripheral neuropathy and localised neuropathies and central nervous system diseases, such as Alzheimer's, Parkinson's disease, Huntington's disease, amyotrophic lateral sclerosis, and Shy-Drager Syndrome. Other uses include the utilisation of the activities such as: Immune system suppression, Activin/inhibin activity, chemotactic/chemokinetic activity, haemostatic and thrombolytic activity, cancer diagnosis and therapy, drug screening, assays for receptor activity, arthritis and inflammation, leukaemias and C.N.S disorders. Note: The sequence data for this patient did not form part of the printed specification
 XX Sequence 513 AA;
 Query Match Score 410.5; DB 4; Length 513;
 Best Local Similarity 71.6%; Pred. No. 3.2e-41;
 Matches 73; Conservative 15; Mismatches 13; Indels 1; Gaps 1;
 CC
 Qy 1 KKTTPRANQNMSTAYLITTAIMASPEKRLTQAQYEMWVNPYFRDKGDNSSSAGWK 60
 Db 100 RKGGSRANQNOSTAELISQASAEPEKRLTQAQYEMWRTVPPFKDGDNSSSAGWK 159
 XX
 RESULT 15
 AAY96447
 ID AAY96447 standard; protein; 501 AA.
 XX
 AC AAY96447;
 XX DT 12-SEP-2000 (first entry)
 DE Forkhead transcription factor Afxf1-501.
 XX
 KW Protein kinase B; PKB; substrate; phosphorylation; forkhead; Afxf;
 KW transcription factor; insulin; suppressor.
 XX
 OS Homo sapiens.
 XX
 PN WO200031291-A1.
 XX
 PN WD200031291-A1.
 XX
 PD 02-JUN-2000.
 XX
 PF 16-NOV-1999; 99WO-SE002095.
 XX
 PR 19-NOV-1998; 98SE-0003958.

XXX (PHAA) PHARMACIA & UPGHORN AB.
XXX James S., Dahlman-Wright K., Lake S., Butcher S., Climent I.

WPI; 2000-400095/34.

Method for screening substances which modulate Protein kinase B activity uses substrate Peptides.

Claim 5: Page 14-16: 41pp: English

This is the forkhead transcription factor, Afx residues 1-501. Afx is phosphorylated by protein kinase B (PKB), and this is enhanced when cells are stimulated with insulin or IGF-1. PKB therefore demonstrates a novel regulation of transcription. AFX-mediated stimulation of IRE driven reporter genes is suppressed by overexpression of PKB and this is further suppressed by insulin. This mechanism can be used to identify compounds for treatment of metabolic disease. Substrate Peptides (AY96441-46) derived from Afx, and forkhead proteins FKHR and FKHL1 (AY96448-49) can be used to discriminate between the effects of compounds which mediate insulin action through transcription via forkhead transcription factor family members from those which modulate activity of enzymes involved in metabolism by phosphorylation. The claimed method comprises screening for substances which are activators, inhibitors or binders of PKB by the use of 2 sequences chosen from AAY96441 and AAY96442, AAY96643 and AAY96644, and AY96645 and AAY96646.

Sequence 501 33:

Query Match	73.2%	Score 406.5 ; DB 3 ; Length 501;
Best Local Similarity	70.6%	Pred. No. 9.6e-41;
Matches	72; Conservative	Mismatches 15; Indels 1; Gaps 1;
Y	1 KRTTTERNAWRGMSYALITTAIASPERKLTLAQAYENWVNQVYERDKGDSNSAGWK 60 88 RKGGGRNARWNQSTAFISIAESAPERKLTLAQAYENWVNRTVYERDKGDSNSAGWK 147	
b	61 NSIRHNLSLHSRFRMILQNEGAGKSSAWNVPD-AKPGMNPR 101 148 NSIRHNLSLISKFKVHEATGKSSAWNVPKGTSGKAPR 149	
Y		
b		

search completed: July 23, 2004, 10:24:14

This Page Blank (Up to)

Copyright (C) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: July 23, 2004, 10:23:16 ; Search time 18 Seconds
(without alignments)
295.415 Million cell updates/sec

Title: US-09-044-353A-54

Perfect score: 555

Sequence: 1 RKTTRRNANGNMSTAYELIT.....SSWWVWINPDAKPGMNNPRRTR 103

Scoring table: BLOSUM62 , Gapext 0.5

Searched: 389414 seqs, 51625971 residues

Total number of hits satisfying chosen parameters: 389414

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%

Listing first 45 summaries

Database : Issued Patents AA:
 1: /cgn2_6:/ptodata/2/iaa/5a_COMB.pep:
 2: /cgn2_6:/ptodata/2/iaa/5b_COMB.pep:
 3: /cgn2_6:/ptodata/2/iaa/6a_COMB.pep:
 4: /cgn2_6:/ptodata/2/iaa/6b_COMB.pep:
 5: /cgn2_6:/ptodata/2/iaa/PCTUS_COMB.pep:
 6: /cgn2_6:/ptodata/2/iaa/backfiles1.pep:
 *

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

SEQ ID NO.	Score	Query Match	Length	DB ID	Description
1	555	100.0	103	US-08-857-076-54	Sequence 54, Appli
2	555	100.0	106	US-08-857-076-37	Sequence 37, Appli
3	555	100.0	109	US-08-857-076-56	Sequence 56, Appli
4	555	100.0	510	US-08-857-076-45	Sequence 45, Appli
5	427.5	77.0	655	US-08-857-076-57	Sequence 57, Appli
6	410.5	74.0	358	US-09-845-629-23	Sequence 23, Appli
7	406.5	73.2	501	US-08-857-076-102	Sequence 102, Appli
8	390.5	70.4	99	US-09-083-351-21	Sequence 21, Appli
9	390.5	70.4	99	US-09-083-352-21	Sequence 21, Appli
10	378.5	68.2	509	US-08-857-076-46	Sequence 46, Appli
11	378.5	68.2	635	US-08-857-076-101	Sequence 101, Appli
12	378	68.1	98	US-08-857-076-58	Sequence 58, Appli
13	188.5	34.0	106	US-09-083-351-16	Sequence 16, Appli
14	188.5	34.0	106	US-09-083-352-16	Sequence 16, Appli
15	185.5	33.4	106	US-09-083-351-17	Sequence 17, Appli
16	185.5	33.4	106	US-09-083-352-17	Sequence 17, Appli
17	185.5	33.4	480	US-07-882-292-2	Sequence 2, Appli
18	185.5	33.4	480	US-08-331-64-2	Sequence 2, Appli
19	185.5	33.4	480	PCT-US93-0102-2	Sequence 2, Appli
20	182.5	32.9	106	US-09-083-351-5	Sequence 5, Appli
21	182.5	32.9	106	US-09-083-351-12	Sequence 12, Appli
22	182.5	32.9	106	US-09-083-351-13	Sequence 13, Appli
23	182.5	32.9	106	US-09-083-352-5	Sequence 5, Appli
24	182.5	32.9	106	US-09-083-352-12	Sequence 12, Appli
25	182.5	32.9	106	US-09-083-352-13	Sequence 13, Appli
26	182.5	32.9	553	US-09-083-351-2	Sequence 2, Appli
27	182.5	32.9	553	US-09-083-352-2	Sequence 2, Appli

ALIGNMENTS

RESULT 1
US-08-857-076-54
; Sequence 54, Application US/08857076C
; Patent No. 6225120
; GENERAL INFORMATION:
; APPLICANT: Ruvkun, Gary
; APPLICANT: Kimura, Routharou
; APPLICANT: Patterson, Garth
; APPLICANT: Ogg, Scott
; APPLICANT: Paradis, Suzanne
; APPLICANT: Tissenbaum, Heidi
; APPLICANT: Morris, Jason
; APPLICANT: Kowek, Allison
; TITLE OF INVENTION: THERAPEUTIC AND DIAGNOSTIC TOOLS FOR IMPAIRED GLUCOSE TOLERANCE CONDITIONS
; FILE REFERENCE: 00186/351001
; CURRENT APPLICATION NUMBER: US/08/857,076C
; CURRENT FILING DATE: 1997-05-15
; NUMBER OF SEQ ID NOS: 114
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 54
; LENGTH: 103
; TYPE: PRT
; ORGANISM: Caenorhabditis elegans

Query Match 100.0%; Score 555, DB 3; Length 103;
Best Local Similarity 100.0%; Pred. No. 8.8e-62;
Matches 103; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 KKTTRRNANGNMSTAYELITAIMASPEKRITLAQYEWNNQNVPYFRDKGDSNSAGWK 60
Db 1 KKTTRRNANGNMSTAYELITAIMASPEKRITLAQYEWNNQNVPYFRDKGDSNSAGWK 60

RESULT 2
US-08-857-076-37
; Sequence 37, Application US/08857076C
; Patent No. 6225120
; GENERAL INFORMATION:
; APPLICANT: Ruvkun, Gary
; APPLICANT: Kimura, Routharou
; APPLICANT: Patterson, Garth
; APPLICANT: Ogg, Scott
; APPLICANT: Paradis, Suzanne
; APPLICANT: Tissenbaum, Heidi

APPLICANT: Morris, Jason
; APPLICANT: Kowek, Allison
; TITLE OF INVENTION: THERAPEUTIC AND DIAGNOSTIC TOOLS FOR
; IMPAIRED GLUCOSE TOLERANCE CONDITIONS
; FILE REFERENCE: 00786/351001
; CURRENT APPLICATION NUMBER: US/08/857,076C
; CURRENT FILING DATE: 1997-05-15
; NUMBER OF SEQ ID NOS: 114
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO: 37
; LENGTH: 106
; TYPE: PRT
; ORGANISM: Caenorhabditis elegans
; US-08-857-076-37

Query Match 100.0%; Score 555; DB 3; Length 106;
Best Local Similarity 100.0%; Pred. No. 9.1e-62;
Matches 103; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 KKTTRRNAGNMSYAELTTAIMASPEKRLLAQYEWVQNYFDRKGDSNSAGWK 60
Db 1 KKTTRRNAGNMSYAELTTAIMASPEKRLLAQYEWVQNYFDRKGDSNSAGWK 60

Qy 61 NSIRHNLSLHSRMRIONEGAKGKSSWWVNPDAKPGMNPRTTR 103
Db 61 NSIRHNLSLHSRMRIONEGAKGKSSWWVNPDAKPGMNPRTTR 103

RESULT 3
US-08-857-076-56
Sequence 56, Application US/08857076C
; Patent No. 6225120
; GENERAL INFORMATION:
; APPLICANT: Ruvkun, Gary
; APPLICANT: Kimura, Koutarou
; APPLICANT: Patterson, Garth
; APPLICANT: Ogg, Scott
; APPLICANT: Paradis, Suzanne
; APPLICANT: Tissenbaum, Heidi
; APPLICANT: Morris, Jason
; APPLICANT: Kowek, Allison
; TITLE OF INVENTION: THERAPEUTIC AND DIAGNOSTIC TOOLS FOR
; IMPAIRED GLUCOSE TOLERANCE CONDITIONS
; FILE REFERENCE: 00786/351001
; CURRENT APPLICATION NUMBER: US/08/857,076C
; CURRENT FILING DATE: 1997-05-15
; NUMBER OF SEQ ID NOS: 114
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO: 56
; LENGTH: 109
; TYPE: PRT
; ORGANISM: Caenorhabditis elegans
; US-08-857-076-56

Query Match 100.0%; Score 555; DB 3; Length 109;
Best Local Similarity 100.0%; Pred. No. 9.5e-62;
Matches 103; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 KKTTRRNAGNMSYAELTTAIMASPEKRLLAQYEWVQNYFDRKGDSNSAGWK 60
Db 7 KKTTRRNAGNMSYAELTTAIMASPEKRLLAQYEWVQNYFDRKGDSNSAGWK 66

Qy 61 NSIRHNLSLHSRMRIONEGAKGKSSWWVNPDAKPGMNPRTTR 103
Db 67 NSIRHNLSLHSRMRIONEGAKGKSSWWVNPDAKPGMNPRTTR 109

RESULT 4
US-08-857-076-45
Sequence 45, Application US/08857076C
; Patent No. 6225120
; GENERAL INFORMATION:
; APPLICANT: Ruvkun, Gary

RESULT 6

US-09-645-629-23 ; Sequence 23, Application US/09645629
; Patent No. 6477515
; GENERAL INFORMATION:
; APPLICANT: Climent Johansson, Isabel
; APPLICANT: Dahlman-Wright, Karin
; APPLICANT: Lake, Staffan
; APPLICANT: Wasserman, Wyeth
; TITLE OF INVENTION: NOVEL RESPONSE ELEMENT
; FILE REFERENCE: 13425-032001
; CURRENT FILING DATE: 2000-08-24
; PRIOR APPLICATION NUMBER: SE 9904269-9
; PRIORITY FILING DATE: 1999-11-25
; PRIOR APPLICATION NUMBER: US 60/151,867
; PRIORITY FILING DATE: 1999-08-31
; PRIOR APPLICATION NUMBER: SE 9903009-0
; PRIOR FILING DATE: 1999-08-26
; NUMBER OF SEQ ID NOS: 34
; SEQ ID NO: 23
; LENGTH: 358
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: synthetically generated GST-AfxDBD construct
; US-09-645-629-23

Query Match 74.0%; Score 410.5; DB 4; Length 358;
Best Local Similarity 71.6%; Pred. No. 5 4e-43; Indels 1; Gaps 1;
Matches 73; Conservative 15; Mismatches 13; Indels 1; Gaps 1;

Qy 1 KKTTRRNANGMSYALITTAIMASPEKRLTIAQTYEWMVNVYPFRDKGDSNSAGWK 60
Db 239 RXGGSRNANGMSYALISQAESAEKRLTIAQTYEWMVNVYPFRDKGDSNSAGWK 298

Qy 61 NSIREHNLSLHESRFMRILONEGAKSSWAVINPD-AKGGMNPRR 101
Db 299 NSIREHNLSLHESRFKTVNEATGKSSWMLNPEGGKSCKSPRR 340

RESULT 7
US-08-857-076-102
; Sequence 102, Application US/08857076C
; Patent No. 6225120
; GENERAL INFORMATION:
; APPLICANT: Ruvkun, Gary
; APPLICANT: Kimura, Koitarou
; APPLICANT: Patterson, Garth
; APPLICANT: Cogg, Scott
; APPLICANT: Paradis, Suzanne
; APPLICANT: Tissenbaum, Heidi
; APPLICANT: Morris, Jason
; APPLICANT: Kowee, Allison
; TITLE OF INVENTION: THERAPEUTIC AND DIAGNOSTIC TOOLS FOR
; IMPAIRED GLUCOSE TOLERANCE CONDITIONS
; FILE REFERENCE: 007B6/351001
; CURRENT APPLICATION NUMBER: US/08/857,076C
; CURRENT FILING DATE: 1997-05-15
; NUMBER OF SEQ ID NOS: 114
; SEQ ID NO: 102
; LENGTH: 501
; TYPE: PRT
; ORGANISM: Homo sapiens
; US-08-857-076-102

Query Match 73.2%; Score 406.5; DB 3; Length 501;
Best Local Similarity 70.6%; Pred. No. 2.7e-42;
Matches 72; Conservative 15; Mismatches 14; Indels 1; Gaps 1;

Qy 1 KKTTRRNANGMSYALITTAIMASPEKRLTIAQTYEWMVNVYPFRDKGDSNSAGWK 60
Db 1 NSIREHNLSLHESRFKTVNEATGKSSWMLNPEGGKSCKSPRR 93

RESULT 8
US-09-083-351-21
; Sequence 21, Application US/09083351
; Patent No. 6087107
; GENERAL INFORMATION:
; APPLICANT: Sheffield, Val C.
; APPLICANT: Alward, Wallace L.M.
; APPLICANT: Stone, Edwin M.
; APPLICANT: Nishimura, Darryl
; APPLICANT: Patil, Shiva
; TITLE OF INVENTION: THERAPEUTICS AND DIAGNOSTICS FOR
; CONGENITAL HEART DISEASE BASED ON A NOVEL HUMAN
; TRANSCRIPTION FACTOR
; NUMBER OF SEQUENCES: 22
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: FOLEY, HORG & ELIOT LLP
; STREET: One Post Office Square
; CITY: Boston
; STATE: MA
; ZIP: 02109-2170
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/083,351
; FILING DATE: 22-MAY-1998
; CLASSIFICATION: 514
; ATTORNEY/AGENT INFORMATION:
; NAME: Arnold, Beth E.
; REGISTRATION NUMBER: 35,430
; REFERENCE/DOCKET NUMBER: UIA-029.02
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 617-832-1000
; TELEFAX: 617-832-7000
; INFORMATION FOR SEQ ID NO: 21:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 99 amino acids
; TYPE: amino acid
; STRANDEDNESS:
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-09-083-351-21

Query Match 70.4%; Score 390.5; DB 3; Length 99;
Best Local Similarity 74.2%; Pred. No. 2.9e-41;
Matches 69; Conservative 15; Mismatches 8; Indels 1; Gaps 1;

Qy 10 WGMSYALITTAIMASPEKRLTIAQTYEWMVNVYPFRDKGDSNSAGWKNSIRHNLSL 69
Db 1 WGNSTADLITKAESSAEKRLTISQYEWMMVKSVPFRDKGDSNSAGWKNSIRHNLSL 60

RESULT 9
US-09-083-352-21
; Sequence 21, Application US/09083352
; Patent No. 6307450
; GENERAL INFORMATION:
; APPLICANT: Sheffield, Val C.
; APPLICANT: Alward, Wallace L.M.
; APPLICANT: Nishimura, Darryl
; APPLICANT: Patil, Shiva
; TITLE OF INVENTION: THERAPEUTICS AND DIAGNOSTICS FOR
; CONGENITAL HEART DISEASE BASED ON A NOVEL HUMAN
; TRANSCRIPTION FACTOR
; NUMBER OF SEQUENCES: 22
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: FOLEY, HORG & ELIOT LLP
; STREET: One Post Office Square
; CITY: Boston
; STATE: MA
; ZIP: 02109-2170
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/083,351
; FILING DATE: 22-MAY-1998
; CLASSIFICATION: 514
; ATTORNEY/AGENT INFORMATION:
; NAME: Arnold, Beth E.
; REGISTRATION NUMBER: 35,430
; REFERENCE/DOCKET NUMBER: UIA-029.02
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 617-832-1000
; TELEFAX: 617-832-7000
; INFORMATION FOR SEQ ID NO: 21:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 99 amino acids
; TYPE: amino acid
; STRANDEDNESS:
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-09-083-351-21

Query Match 70.4%; Score 390.5; DB 3; Length 99;
Best Local Similarity 74.2%; Pred. No. 2.9e-41;
Matches 69; Conservative 15; Mismatches 8; Indels 1; Gaps 1;

Qy 70 HSREMRJLONEGAKSSWAVINPD-AKGGMNPRR 101
Db 1 WGNSTADLITKAESSAEKRLTISQYEWMMVKSVPFRDKGDSNSAGWKNSIRHNLSL 60

Qy 70 HSREMRJLONEGAKSSWAVINPD-AKGGMNPRR 101
Db 61 HSFKFIRQNNEGTKSSWMLNPEGGSKSPRR 93

CURRENT FILING DATE: 1997-05-15
 SOFTWARE: FastSEQ for Windows Version 4.0
 SEQ ID NO: 58
 LENGTH: 98
 TYPE: PRT
 ORGANISM: Caenorhabditis elegans
 US-08-857-076-58

Query Match 68.1%; Score 378; DB 3; Length 98;
 Best Local Similarity 66.7%; Pred. No. 1.le-39;
 Matches 64; Conservative 18; Mismatches 0; Gaps 0;
 Qy 8 NAWGNMMSYELITTAIASPERKLTLAQVYEMVNQVNPYFRDGDSNSAGWKNSIRHNL 67
 Db 3 NPWGEESYSDIIAKALESPDGRKLNEYQWFSNDNIPPFGERSSPEAAAGWKNSIRHNL 62
 Qy 68 SLHSRFMRITQNEGAKSSWNVINPDAKPGMNPRTR 103
 Db 63 SLHSRFMRITQNEGAKSSWNVINPDAKPGMNPRTR 98

RESULT 13

US-09-083-351-16

Sequence 16, Application US/09083351
 Patent No. 6087107
 GENERAL INFORMATION:
 APPLICANT: Sheffield, Val C.
 APPLICANT: Alward, Wallace L.M.
 APPLICANT: Stone, Edwin M.
 APPLICANT: Nishimura, Darryl
 APPLICANT: Patil, Shiva
 TITLE OF INVENTION: THERAPEUTICS AND DIAGNOSTICS FOR CONGENITAL HEART DISEASE BASED ON A NOVEL HUMAN TRANSCRIPTION FACTOR
 NUMBER OF SEQUENCES: 22
 CORRESPONDENCE ADDRESS:
 ADDRESSEE: FOLEY, HOAG & ELIOT LLP
 STREET: One Post Office Square
 CITY: Boston
 STATE: MA
 COUNTRY: USA
 ZIP: 02109-2170
 COMPUTER READABLE FORM:
 MEDIUM TYPE: Floppy disk
 COMPUTER: IBM PC compatible
 OPERATING SYSTEM: PC-DOS/MS-DOS
 SOFTWARE: Patent In Release #1.0, Version #1.30
 CURRENT APPLICATION DATA:
 APPLICATION NUMBER: US/09/083,352
 FILING DATE: 22-MAY-1998
 CLASSIFICATION: 514
 ATTORNEY/AGENT INFORMATION:
 NAME: Arnold, Beth E.
 REGISTRATION NUMBER: JIA-029.01
 REFERENCE/DOCKET NUMBER: JIA-029.01
 TELECOMMUNICATION INFORMATION:
 TELEPHONE: 617-832-1000
 TELEFAX: 617-832-7000
 INFORMATION FOR SEQ ID NO: 16:
 SEQUENCE CHARACTERISTICS:
 LENGTH: 106 amino acids
 TYPE: amino acid
 STRANDEDNESS:
 TOPOLOGY: linear
 MOLECULE TYPE: protein

US-09-083-352-16

Query Match 34.0%; Score 188.5; DB 3; Length 106;
 Best Local Similarity 43.9%; Pred. No. 5.5e-16;
 Matches 36; Conservative 20; Mismatches 19; Gaps 2;
 Qy 14 SYAELTTAIASPERKLTLAQVYEMVNQVNPYFRDGDSNSAGWKNSIRHNLHSRF 73
 Db 10 SYNALIMMA-TQSPKRLTINGIYEFLIMKNPYRE---NKOGWQNSTRHNLSLNKCF 64
 Qy 74 MRION-EGAGKSSWVINPDA 93
 Db 65 VKVPRHYDDPGKGNWMLDSS 86

RESULT 15

US-09-083-351-17

Sequence 17, Application US/09083351
 Patent No. 6087107
 GENERAL INFORMATION:
 APPLICANT: Sheffield, Val C.
 APPLICANT: Alward, Wallace L.M.
 TITLE OF INVENTION: GLAUCOMA THERAPEUTICS AND DIAGNOSTICS BASED ON A NOVEL HUMAN TRANSCRIPTION FACTOR
 NUMBER OF SEQUENCES: 22
 CORRESPONDENCE ADDRESS:
 ADDRESSEE: FOLEY, HOAG & ELIOT LLP
 STREET: One Post Office Square
 CITY: Boston
 STATE: MA
 COUNTRY: USA
 ZIP: 02109-2170
 COMPUTER READABLE FORM:
 MEDIUM TYPE: Floppy disk
 COMPUTER: IBM PC compatible
 OPERATING SYSTEM: PC-DOS/MS-DOS
 SOFTWARE: Patent In Release #1.0, Version #1.30
 CURRENT APPLICATION DATA:
 APPLICATION NUMBER: US/09/083,352
 FILING DATE: 22-MAY-1998
 CLASSIFICATION: 514
 ATTORNEY/AGENT INFORMATION:
 NAME: Arnold, Beth E.
 REGISTRATION NUMBER: JIA-029.01
 REFERENCE/DOCKET NUMBER: JIA-029.01
 TELECOMMUNICATION INFORMATION:
 TELEPHONE: 617-832-1000
 TELEFAX: 617-832-7000
 INFORMATION FOR SEQ ID NO: 16:
 SEQUENCE CHARACTERISTICS:
 LENGTH: 106 amino acids
 TYPE: amino acid
 STRANDEDNESS:
 TOPOLOGY: linear
 MOLECULE TYPE: protein

US-09-083-352-16

i APPLICANT: Stone, Edwin M.
 i APPLICANT: Nishimura, Darryl
 i APPLICANT: Pacil, Shiva
 i TITLE OF INVENTION: THERAPEUTICS AND DIAGNOSTICS FOR
 i CONGENITAL HEART DISEASE BASED ON A NOVEL HUMAN
 i TITLE OF INVENTION: TRANSCRIPTION FACTOR
 i NUMBER OF SEQUENCES: 22
 i CORRESPONDENCE ADDRESS:
 i ADDRESSEE: FOLLY, ROG & ELIOT LLP
 i STREET: One Post Office Square
 i CITY: Boston
 i STATE: MA
 i COUNTRY: USA
 i ZIP: 02109-2170
 i COMPUTER READABLE FORM:
 i MEDIUM TYPE: Floppy disk
 i COMPUTER: IBM PC compatible
 i OPERATING SYSTEM: PC-DOS/MS-DOS
 i SOFTWARE: PatentIn Release #1.0, Version #1.3.0
 i CURRENT APPLICATION DATA:
 i APPLICATION NUMBER: US/09/083,351
 i FILING DATE: 22-MAY-1998
 i CLASSIFICATION: 514
 i ATTORNEY/AGENT INFORMATION:
 i NAME: Arnold, Beth E.
 i REFERENCE/DOCKET NUMBER: 35,430
 i TELECOMMUNICATION INFORMATION:
 i TELEPHONE: 617-332-1000
 i TELEFAX: 617-832-7000
 i INFORMATION FOR SEQ ID NO: 17:
 i SEQUENCE CHARACTERISTICS:
 i LENGTH: 106 amino acids
 i TYPE: amino acid
 i STRANDEDNESS:
 i TOPOLOGY: linear
 i MOLECULE TYPE: protein
 US-09-083-351-17

Query Match 33.4%; Score 185; DB 3; Length 106;
 Best Local Similarity 42.7%; Pred. No. 1.3e-15;
 Matches 35; Conservative 21; Mismatches 19; Indels 7; Gaps 2;
 Qy 14 SYAELTTIAINASPERKLTIAQVYEMVNQNPYFRDKGDSNSAGMKNSIRHNLSLHSRF 73
 Db 10 SYNALIMMAMRQSPEKRLTNGIYEFIMKNPYYRE-----NKQGMQNSIRHNLSLNKCF 64
 Qy 74 MRION-EGAKSSMWVNPDA 93
 Db 65 VKVPRHYDDPGKGNYWWLDPSS 86

Search completed: July 23, 2004, 10:26:23
 Job time : 18 secs

Copyright (c) 1993 - 2004 Compugen Ltd.

Gencore version 5.1.6

OM protein - protein search, using SW mode.

Run on: July 23, 2004, 10:25:31 ; Search time 47 Seconds
 (without alignments)
 686.274 Million cell updates/sec

Title: US-09-844-353A-54

Scoring table: BLOSUM62

Gapext 0.5

Perfect score: 555

Sequence: 1 KKTITRRAWNMSYALIT.....SSWWVTPDAKPGNMNPRTR 103

Searched: 1288442 seqs, 313154207 residues

Total number of hits satisfying chosen parameters: 1288442

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
 Maximum Match 10%

Listing first 45 summaries

Database : Published Applications AA:
 1: /cgm2_6/ptodata/2/pubpa/US07_PUBCOMB.pep:*

2: /cgm2_6/ptodata/2/pubpa/BCT_NEW_PUB.pep:*

3: /cgm2_6/ptodata/2/pubpa/US06_NEW_PUB.pep:*

4: /cgm2_6/ptodata/2/pubpa/US05_PUBCOMB.pep:*

5: /cgm2_6/ptodata/2/pubpa/US07_NEW_PUB.pep:*

6: /cgm2_6/ptodata/2/pubpa/PCTUS_PUBCOMB.pep:*

7: /cgm2_6/ptodata/2/pubpa/US08_NEW_PUB.pep:*

8: /cgm2_6/ptodata/2/pubpa/US09_NEW_PUB.pep:*

9: /cgm2_6/ptodata/2/pubpa/US09A_PUBCOMB.pep:*

10: /cgm2_6/ptodata/2/pubpa/US09B_PUBCOMB.pep:*

11: /cgm2_6/ptodata/2/pubpa/US09C_PUBCOMB.pep:*

12: /cgm2_6/ptodata/2/pubpa/US09_NEW_PUB.pep:*

13: /cgm2_6/ptodata/2/pubpa/US10A_PUBCOMB.pep:*

14: /cgm2_6/ptodata/2/pubpa/US10B_PUBCOMB.pep:*

15: /cgm2_6/ptodata/2/pubpa/US10C_PUBCOMB.pep:*

16: /cgm2_6/ptodata/2/pubpa/US10_NEW_PUB.pep:*

17: /cgm2_6/ptodata/2/pubpa/US60_NEW_PUB.pep:*

18: /cgm2_6/ptodata/2/pubpa/US60_PUBCOMB.pep:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query	Match	Length	DB ID	Description
1	555	100.0	103	9	US-09-205-658-54	Sequence 57, Appl
2	555	100.0	103	9	US-09-244-353A-54	Sequence 53, Appl
3	555	100.0	103	10	US-09-963-693-54	Sequence 54, Appl
4	555	100.0	106	9	US-09-205-658-37	Sequence 54, Appl
5	555	100.0	106	9	US-09-644-353A-37	Sequence 37, Appl
6	555	100.0	106	10	US-09-963-693-37	Sequence 37, Appl
7	555	100.0	109	9	US-09-205-658-56	Sequence 56, Appl
8	555	100.0	109	9	US-09-844-353A-56	Sequence 56, Appl
9	555	100.0	109	10	US-09-963-693-56	Sequence 56, Appl
10	555	100.0	101	9	US-09-205-558-45	Sequence 45, Appl
11	555	100.0	510	9	US-09-844-353A-45	Sequence 45, Appl
12	555	100.0	510	10	US-09-963-693-45	Sequence 45, Appl
13	440.5	79.4	673	16	US-10-701-490-6	Sequence 6, Appl
14	427.5	77.0	655	9	US-09-205-558-57	Sequence 57, Appl
15	427.5	77.0	655	9	US-09-844-353A-57	Sequence 57, Appl

ALIGNMENTS

SEQ ID NO	TYPE	ORGANISM
54	PRT	Caenorhabditis elegans
54	PRT	Caenorhabditis elegans

RESULT 1
 US-09-205-658-54
 ; Sequence 54, Application US/09205658
 ; GENERAL INFORMATION;
 ; PATENT NO. US2010029617A1
 ; APPLICANT: Ruvkin, Gary
 ; APPLICANT: Ogg, Scott
 ; TITLE OF INVENTION: THERAPEUTIC AND DIAGNOSTIC TOOLS FOR IMPAIRED GLUCOSE TOLERANCE CONDITIONS
 ; TITLE OF INVENTION: IMPAIRED GLUCOSE TOLERANCE CONDITIONS
 ; FILE REFERENCE: 00786/351004
 ; CURRENT APPLICATION NUMBER: US/09/205,658
 ; CURRENT FILING DATE: 1998-12-03
 ; EARLIER APPLICATION NUMBER: US/09/205,658
 ; EARLIER FILING DATE: 1998-05-15
 ; EARLIER APPLICATION NUMBER: US/09/205,657
 ; EARLIER FILING DATE: 1998-05-15
 ; NUMBER OF SEQ ID NOS: 328
 ; SOFTWARE: FastSEQ for Windows Version 4.0
 ; SEQ ID NO 54
 ; LENGTH: 103
 ; TYPE: PRT
 ; ORGANISM: Caenorhabditis elegans

QY 1 KKTTTRRNAGNMNSYALITTAIASPERLTLAQYENVNQVNPYFRXGDSNNSAGWK 60
 ; Query Match 100.0%; Score 555; DB 9; Length 103;
 ; Best Local Similarity 100.0%; Mismatches 0; Indels 0; Gaps 0;
 ; Matches 103; Conservative 0;
 ; Best Local Similarity 100.0%; Mismatches 0; Indels 0; Gaps 0;

Db 1 KKTTTRRNAGNMNSYALITTAIASPERLTLAQYENVNQVNPYFRXGDSNNSAGWK 60
 ; Query Match 100.0%; Score 555; DB 9; Length 103;
 ; Best Local Similarity 100.0%; Mismatches 0; Indels 0; Gaps 0;

QY 1 NSIRENLISLSERMRNQNEGAKSSWWVINDAKPGMNPRTR 103
 ; Query Match 100.0%; Score 555; DB 9; Length 103;
 ; Best Local Similarity 100.0%; Mismatches 0; Indels 0; Gaps 0;

Db 1 NSIRENLISLSERMRNQNEGAKSSWWVINDAKPGMNPRTR 103
 ; Query Match 100.0%; Score 555; DB 9; Length 103;
 ; Best Local Similarity 100.0%; Mismatches 0; Indels 0; Gaps 0;

QY 1 NSIRHNLISLSERMRNQNEGAKSSWWVINDAKPGMNPRTR 103
 ; Query Match 100.0%; Score 555; DB 9; Length 103;
 ; Best Local Similarity 100.0%; Mismatches 0; Indels 0; Gaps 0;

RESULT 2
 US-09-844-353A-54
 ; Sequence 54, Application US/09844353A
 ; Patent No. US2002003758A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Ruvkun, Gary
 ; APPLICANT: Kimura, Koutarou
 ; APPLICANT: Patterson, Garth
 ; APPLICANT: Ogg, Scott
 ; APPLICANT: Paradis, Suzanne
 ; APPLICANT: Tiessenbaum, Heidi
 ; APPLICANT: Morris, Jason
 ; APPLICANT: Kowek, Allison
 ; TITLE OF INVENTION: THERAPEUTIC AND DIAGNOSTIC TOOLS FOR IMPAIRED GLUCOSE TOLERANCE CONDITIONS
 ; FILE REFERENCE: 0786/35105
 ; CURRENT APPLICATION NUMBER: US/09/844,353A
 ; PRIOR APPLICATION NUMBER: US 2001-04-27
 ; PRIOR FILING DATE: 2001-04-27
 ; NUMBER OF SEQ ID NOS: 114
 ; SOFTWARE: FastSEQ for Windows Version 4.0
 ; SEQ ID NO: 54
 ; LENGTH: 103
 ; TYPE: PRT
 ; ORGANISM: Caenorhabditis elegans
 US-09-844-353A-54

Query Match 100.0%; Score 555; DB 9; Length 103;
 Best Local Similarity 100.0%; Pred. No. 3.4e-58;
 Matches 103; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 KKTTRRNAGNMSTAELTTTAAASPEKRLLAQYEWVNQVYFRDKGDSNSAGWK 60
 Db 1 KKTTRRNAGNMSTAELTTTAAASPEKRLLAQYEWVNQVYFRDKGDSNSAGWK 60

Qy 61 NSIRHNLSLHSRMRIONEGAKSSWWVNPDAKPGMNPRTTR 103
 Db 61 NSIRHNLSLHSRMRIONEGAKSSWWVNPDAKPGMNPRTTR 103

RESULT 3
 US-09-963-693-54
 ; Sequence 54, Application US/09963693
 ; Publication No. US20030181364A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Ruvkun, Gary
 ; APPLICANT: Ogg, Scott
 ; TITLE OF INVENTION: THERAPEUTIC AND DIAGNOSTIC TOOLS FOR IMPAIRED GLUCOSE TOLERANCE CONDITIONS
 ; FILE REFERENCE: 0786/351004
 ; CURRENT APPLICATION NUMBER: US/09/963,693
 ; PRIOR APPLICATION NUMBER: US/09/205,658
 ; PRIOR APPLICATION NUMBER: 08/857,076
 ; PRIOR FILING DATE: 1997-05-15
 ; PRIOR FILING DATE: 1997-07-07
 ; PRIOR APPLICATION NUMBER: US/09/10080
 ; SEQ ID NO: 54
 ; LENGTH: 103
 ; TYPE: PRT
 ; ORGANISM: Caenorhabditis elegans
 US-09-963-693-54

Query Match 100.0%; Score 555; DB 10; Length 103;
 Best Local Similarity 100.0%; Pred. No. 3.4e-58;
 Matches 103; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 KKTTRRNAGNMSTAELTTTAAASPEKRLLAQYEWVNQVYFRDKGDSNSAGWK 60
 Db 1 KKTTRRNAGNMSTAELTTTAAASPEKRLLAQYEWVNQVYFRDKGDSNSAGWK 60

Qy 61 NSIRHNLSLHSRMRIONEGAKSSWWVNPDAKPGMNPRTTR 103
 Db 61 NSIRHNLSLHSRMRIONEGAKSSWWVNPDAKPGMNPRTTR 103

RESULT 4
 US-09-205-658-37
 ; Sequence 37, Application US/09205658
 ; Patent No. US20010023617A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Ruvkun, Gary
 ; APPLICANT: Ogg, Scott
 ; TITLE OF INVENTION: THERAPEUTIC AND DIAGNOSTIC TOOLS FOR IMPAIRED GLUCOSE TOLERANCE CONDITIONS
 ; FILE REFERENCE: 00786/351004
 ; CURRENT APPLICATION NUMBER: US/09/205,658
 ; CURRENT FILING DATE: 1998-12-03
 ; EARLIER APPLICATION NUMBER: 08/857,076
 ; EARLIER FILING DATE: 1997-05-15
 ; EARLIER APPLICATION NUMBER: 08/888,534
 ; EARLIER FILING DATE: 1997-07-07
 ; EARLIER APPLICATION NUMBER: US98/10080
 ; EARLIER FILING DATE: 1998-07-07
 ; SEQ ID NO: 328
 ; SOFTWARE: FastSEQ for Windows Version 4.0
 ; NUMBER OF SEQ ID NOS: 328
 ; TYPE: PRT
 ; ORGANISM: Caenorhabditis elegans
 US-09-205-658-37

Query Match 100.0%; Score 555; DB 9; Length 106;
 Best Local Similarity 100.0%; Pred. No. 3.5e-58;
 Matches 103; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 KKTTRRNAGNMSTAELTTTAAASPEKRLLAQYEWVNQVYFRDKGDSNSAGWK 60
 Db 1 KKTTRRNAGNMSTAELTTTAAASPEKRLLAQYEWVNQVYFRDKGDSNSAGWK 60

Qy 61 NSIRHNLSLHSRMRIONEGAKSSWWVNPDAKPGMNPRTTR 103
 Db 61 NSIRHNLSLHSRMRIONEGAKSSWWVNPDAKPGMNPRTTR 103

RESULT 5
 US-09-844-353A-37
 ; Sequence 37, Application US/09844353A
 ; Patent No. US2002003758A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Ruvkun, Gary
 ; APPLICANT: Kimura, Koutarou
 ; APPLICANT: Patterson, Garth
 ; APPLICANT: Ogg, Scott
 ; APPLICANT: Paradis, Suzanne
 ; APPLICANT: Tiessenbaum, Heidi
 ; APPLICANT: Morris, Jason
 ; APPLICANT: Kowek, Allison
 ; TITLE OF INVENTION: THERAPEUTIC AND DIAGNOSTIC TOOLS FOR IMPAIRED GLUCOSE TOLERANCE CONDITIONS
 ; FILE REFERENCE: 00786/351005
 ; CURRENT APPLICATION NUMBER: US/09/844,353A
 ; CURRENT FILING DATE: 2001-04-27
 ; PRIOR APPLICATION NUMBER: 08/857,076
 ; PRIOR FILING DATE: 1997-05-15
 ; PRIOR FILING DATE: 1997-07-07
 ; PRIOR APPLICATION NUMBER: US/09/10080
 ; SEQ ID NO: 54
 ; LENGTH: 103
 ; TYPE: PRT
 ; ORGANISM: Caenorhabditis elegans
 US-09-963-693-54

Query Match 100.0%; Score 555; DB 10; Length 103;
 Best Local Similarity 100.0%; Pred. No. 3.4e-58;
 Matches 103; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

; TYPE: PRT ; ORGANISM: Caenorhabditis elegans

Query Match 100.0%; Score 555; DB 9; Length 106;
Best Local Similarity 100.0%; Pred. No. 3.5e-58; Indels 0; Gaps 0;
Matches 103; Conservative 0; Mismatches 0; SEQ ID NO 56

Qy 1 KKTTRRNAGNMSYAELTTAIMASPEKRILTAQYEMVQNPYFRDKGDSNSAGWK 60
Db 1 KKTTRRNAGNMSYAELTTAIMASPEKRILTAQYEMVQNPYFRDKGDSNSAGWK 60

RESULT 6 US-09-963-693-37

Qy Sequence 37, Application US/0963693
; Publication No. US2003018136441
; GENERAL INFORMATION:
; APPLICANT: Ruvkun, Gary
; Ogg, Scott

; TITLE OF INVENTION: THERAPEUTIC AND DIAGNOSTIC TOOLS FOR
; FILE REFERENCE: 00786/351004

CURRENT APPLICATION NUMBER: US/09/963.693
CURRENT FILING DATE: 2001-09-25

PRIOR APPLICATION NUMBER: US/09/205,658
PRIOR FILING DATE: 1998-12-03

PRIOR APPLICATION NUMBER: 08/857,076
PRIOR FILING DATE: 1997-05-15

PRIOR APPLICATION NUMBER: 08/888,534
PRIOR FILING DATE: 1997-07-07

PRIOR APPLICATION NUMBER: US98/10080
PRIOR FILING DATE: 1998-05-15

NUMBER OF SEQ ID NOS: 328
SOFTWARE: FastSEQ for Windows Version 4.0
SEQ ID NO 37
LENGTH: 106

TYPE: PRT ; ORGANISM: Caenorhabditis elegans

US-09-963-693-37

Query Match 100.0%; Score 555; DB 10; Length 106;
Best Local Similarity 100.0%; Pred. No. 3.5e-58; Indels 0; Gaps 0;
Matches 103; Conservative 0; Mismatches 0; SEQ ID NO 56

Qy 1 KKTTRRNAGNMSYAELTTAIMASPEKRILTAQYEMVQNPYFRDKGDSNSAGWK 60
Db 1 KKTTRRNAGNMSYAELTTAIMASPEKRILTAQYEMVQNPYFRDKGDSNSAGWK 60

RESULT 7 US-09-205-65-56

Qy Sequence 56, Application US/09205658
; Patent No. US2001029617A1
; GENERAL INFORMATION:
; APPLICANT: Ogg, Scott

; TITLE OF INVENTION: THERAPEUTIC AND DIAGNOSTIC TOOLS FOR
; FILE REFERENCE: 00786/351004

CURRENT APPLICATION NUMBER: US/09/205,658
CURRENT FILING DATE: 1998-12-03

EARLIER APPLICATION NUMBER: 08/857,076
EARLIER FILING DATE: 1997-05-15

EARLIER APPLICATION NUMBER: 08/888,534
EARLIER FILING DATE: 1997-07-07

NUMBER OF SEQ ID NOS: 328
SOFTWARE: FastSEQ for Windows Version 4.0

RESULT 8 US-09-844-353A-56

Qy Sequence 56, Application US/09844353A
; Patent No. US20020037585A1
; GENERAL INFORMATION:
; APPLICANT: Ruvkun, Gary
; Kimura, Konärou
; Patterson, Garth
; Ogg, Scott
; Paradis, Suzanne
; Tissenbaum, Heidi
; Morris, Jason
; Kowek, Allison
; TITLE OF INVENTION: THERAPEUTIC AND DIAGNOSTIC TOOLS FOR
; IMPAIRED GLUCOSE TOLERANCE CONDITIONS
; FILE REFERENCE: 00786/351005

CURRENT APPLICATION NUMBER: US/09/844,353A
CURRENT FILING DATE: 2001-04-27

PRIOR APPLICATION NUMBER: US 08/857,076
PRIOR FILING DATE: 1997-05-15
NUMBER OF SEQ ID NOS: 114
SOFTWARE: FastSEQ for Windows Version 4.0
SEQ ID NO 56

TYPE: PRT ; ORGANISM: Caenorhabditis elegans

US-09-844-353A-56

Query Match 100.0%; Score 555; DB 9; Length 109;
Best Local Similarity 100.0%; Pred. No. 3.6e-58; Indels 0; Gaps 0;

Qy 1 KKTTRRNAGNMSYAELTTAIMASPEKRILTAQYEMVQNPYFRDKGDSNSAGWK 60
Db 7 KKTTRRNAGNMSYAELTTAIMASPEKRILTAQYEMVQNPYFRDKGDSNSAGWK 66

RESULT 9 US-09-963-693-56

Qy Sequence 56, Application US/09963693
; Publication No. US20030181364A1
; GENERAL INFORMATION:
; APPLICANT: Ruvkun, Gary
; Ogg, Scott

; TITLE OF INVENTION: THERAPEUTIC AND DIAGNOSTIC TOOLS FOR

RESULT 10
 TITLE OF INVENTION: IMPAIRED GLUCOSE TOLERANCE CONDITIONS
 FILE REFERENCE: 00786/351004
 CURRENT APPLICATION NUMBER: US/09/963,693
 CURRENT FILING DATE: 2001-09-25
 PRIOR APPLICATION NUMBER: US/09/205,658
 PRIOR FILING DATE: 1998-12-03
 PRIOR APPLICATION NUMBER: 08/857,076
 PRIOR FILING DATE: 1997-05-15
 PRIOR APPLICATION NUMBER: 08/888,534
 PRIOR FILING DATE: 1997-07-07
 PRIOR APPLICATION NUMBER: US/98/10080
 PRIOR FILING DATE: 1998-05-15
 NUMBER OF SEQ ID NOS: 328
 SOFTWARE: FastSEQ for Windows Version 4.0
 SEQ ID NO: 56
 LENGTH: 109
 TYPE: PRT
 ORGANISM: Caenorhabditis elegans
 US-09-963-693-56

Query Match 100.0%; Score 555; DB 10; Length 109;
 Best Local Similarity 100.0%; Pred. No. 3.6e-58; Indels 0; Gaps 0;
 Matches 103; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db 7 KKTTRRNAAGNMSYALITTAIMASPEKRILTAQYEWVNQVYFRDKGDSNSAGWK 60
 Db 61 NSIRHNLSLHSRMRQNEGAKSSWVINDAKPGMNPRTTR 103
 Db 67 NSIRHNLSLHSRMRQNEGAKSSWVINDAKPGMNPRTTR 109

RESULT 10
 TITLE OF INVENTION: IMPAIRED GLUCOSE TOLERANCE CONDITIONS
 FILE REFERENCE: 00786/351004
 CURRENT APPLICATION NUMBER: US/09/205,658
 CURRENT FILING DATE: 1998-12-03
 EARLIER APPLICATION NUMBER: 08/857,076
 EARLIER FILING DATE: 1997-05-15
 EARLIER APPLICATION NUMBER: 08/888,534
 EARLIER FILING DATE: 1997-07-07
 EARLIER APPLICATION NUMBER: US/98/10080
 NUMBER OF SEQ ID NOS: 328
 SOFTWARE: FastSEQ for Windows Version 4.0
 SEQ ID NO: 45
 LENGTH: 510
 TYPE: PRT
 ORGANISM: Caenorhabditis elegans
 US-09-205-658-45

Query Match 100.0%; Score 555; DB 9; Length 510;
 Best Local Similarity 100.0%; Pred. No. 2.3e-57; Indels 0; Gaps 0;
 Matches 103; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db 1 KKTTRRNAAGNMSYALITTAIMASPEKRILTAQYEWVNQVYFRDKGDSNSAGWK 60
 Db 135 NSIRHNLSLHSRMRQNEGAKSSWVINDAKPGMNPRTTR 103
 Db 195 NSIRHNLSLHSRMRQNEGAKSSWVINDAKPGMNPRTTR 237

RESULT 12
 TITLE OF INVENTION: IMPAIRED GLUCOSE TOLERANCE CONDITIONS
 FILE REFERENCE: 00786/351004
 CURRENT APPLICATION NUMBER: US/09/205,658
 CURRENT FILING DATE: 2001-09-25
 EARLIER APPLICATION NUMBER: 08/857,076
 EARLIER FILING DATE: 1998-12-03
 EARLIER APPLICATION NUMBER: 08/888,534
 EARLIER FILING DATE: 1997-05-15
 EARLIER APPLICATION NUMBER: US/98/10080
 NUMBER OF SEQ ID NOS: 328
 SOFTWARE: FastSEQ for Windows Version 4.0
 SEQ ID NO: 45
 LENGTH: 510
 TYPE: PRT
 ORGANISM: Caenorhabditis elegans
 US-09-205-658-45

Query Match 100.0%; Score 555; DB 10; Length 510;
 Best Local Similarity 100.0%; Pred. No. 2.3e-57; Indels 0; Gaps 0;
 Matches 103; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db 1 KKTTRRNAAGNMSYALITTAIMASPEKRILTAQYEWVNQVYFRDKGDSNSAGWK 60
 Db 135 KKTTRRNAAGNMSYALITTAIMASPEKRILTAQYEWVNQVYFRDKGDSNSAGWK 194
 Qy 1 KKTTRRNAAGNMSYALITTAIMASPEKRILTAQYEWVNQVYFRDKGDSNSAGWK 60

RESULT 12
 TITLE OF INVENTION: IMPAIRED GLUCOSE TOLERANCE CONDITIONS
 FILE REFERENCE: 00786/351004
 CURRENT APPLICATION NUMBER: US/09/205,658
 CURRENT FILING DATE: 2001-09-25
 EARLIER APPLICATION NUMBER: 08/857,076
 EARLIER FILING DATE: 1998-12-03
 EARLIER APPLICATION NUMBER: 08/888,534
 EARLIER FILING DATE: 1997-05-15
 EARLIER APPLICATION NUMBER: US/98/10080
 NUMBER OF SEQ ID NOS: 328
 SOFTWARE: FastSEQ for Windows Version 4.0
 SEQ ID NO: 45
 LENGTH: 510
 TYPE: PRT
 ORGANISM: Caenorhabditis elegans
 US-09-205-658-45

Query Match 100.0%; Score 555; DB 10; Length 510;
 Best Local Similarity 100.0%; Pred. No. 2.3e-57; Indels 0; Gaps 0;
 Matches 103; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db 1 KKTTRRNAAGNMSYALITTAIMASPEKRILTAQYEWVNQVYFRDKGDSNSAGWK 60
 Db 135 KKTTRRNAAGNMSYALITTAIMASPEKRILTAQYEWVNQVYFRDKGDSNSAGWK 194
 Qy 1 KKTTRRNAAGNMSYALITTAIMASPEKRILTAQYEWVNQVYFRDKGDSNSAGWK 60

RESULT 13
US-10-701-490-6

; Sequence 6, Application US/10701490

; Publication No. US20040106141A1.

; GENERAL INFORMATION:

; APPLICANT: PAUL S. MISCHEL

; APPLICANT: CHARLES L. SAWYERS

; APPLICANT: BRADLEY L. SMITH

; APPLICANT: KATHERINE CROSBY

; TITLE OF INVENTION: METHODS AND MATERIALS FOR EXAMINING PATHWAYS ASSOCIATED WITH GLIOBLASTOMA PROGRESSION

; TITLE OF INVENTION: CURRENT APPLICATION NUMBER: US11/14USU1

; FILE REFERENCE: G&C

; CURRENT FILING DATE: 2003-11-05

; PRIOR FILING DATE: 2002-11-05

; NUMBER OF SEQ ID NOS: 11

; SOFTWARE: FastSEQ for Windows Version 4.0

; SEQ ID NO: 6

; LENGTH: 673

; TYPE: PRT

; ORGANISM: Homo Sapiens

US-10-701-490-6

Query Match 79.4%; Score 440.5; DB 16; Length 673;
Best Local Similarity 77.5%; Pred. No. 1.5e-43;
Matches 79; Conservative 13; Mismatches 9; Indels 1; Gaps 1;

Query 1 KKTTRRNAAGNMSYAELTTAIMASPEKRLLAQVEMVQNPYFRDKGDSNSAGWK 60
Db 148 RACSSRNAGNMSYAELTTAIMASPEKRLLAQVEMVQNPYFRDKGDSNSAGWK 207

Query 61 NSTRHNLSLHSRMRMIONEAGKSSWWVNPD-AKGMMNPRR 101
Db 208 NSTRHNLSLHSRMRMIONEAGKSSWWVNPDGGKSCKSPRR 249

RESULT 14
US-09-205-658-57

; Sequence 57, Application US/09205658

; Patent No. US20010029617A1

; GENERAL INFORMATION:

; APPLICANT: Ruykun, Gary

; APPLICANT: Ogg, Scott

; APPLICANT: Paradis, Suzanne

; APPLICANT: Tissenbaum, Heidi

; APPLICANT: Morris, Jason

; APPLICANT: Kowek, Allison

; TITLE OF INVENTION: THERAPEUTIC AND DIAGNOSTIC TOOLS FOR IMPAIRED GLUCOSE TOLERANCE CONDITIONS

; FILE REFERENCE: 00786/351005

; CURRENT APPLICATION NUMBER: US/09/844-353A

; CURRENT FILING DATE: 2001-04-27

; PRIOR APPLICATION NUMBER: US 08/857,076

; PRIOR FILING DATE: 1997-05-15

; NUMBER OF SEQ ID NOS: 114

; SOFTWARE: FastSEQ for Windows Version 4.0

; SEQ ID NO: 57

; LENGTH: 655

; TYPE: PRT

; ORGANISM: Homo sapiens

US-09-844-353A-57

Query Match 77.0%; Score 427.5; DB 9; Length 655;
Best Local Similarity 73.5%; Pred. No. 5.1e-42;
Matches 75; Conservative 18; Mismatches 8; Indels 1; Gaps 1;

Query 1 KKTTRRNAAGNMSYAELTTAIMASPEKRLLAQVEMVQNPYFRDKGDSNSAGWK 60
Db 151 KSSSRNAGNMSYAELTTAIMASPEKRLLAQVEMVQNPYFRDKGDSNSAGWK 210

Query 61 NSTRHNLSLHSRMRMIONEAGKSSWWVNPD-AKGMMNPRR 101
Db 211 NSTRHNLSLHSRMRMIONEAGKSSWWVNPDGGKSCKSPRR 252

Search completed: July 23, 2004, 10:31:01
Job time : 47 secs

RESULT 15
US-09-844-353A-57

; Sequence 57, Application US/09844353A

; Patent No. US20020031585A1

; GENERAL INFORMATION:

; APPLICANT: Ruykun, Gary

; APPLICANT: Kimura, Koutarou

; APPLICANT: Patterson, Garth

; APPLICANT: Ogg, Scott

; APPLICANT: Paradis, Suzanne

; APPLICANT: Tissenbaum, Heidi

; APPLICANT: Morris, Jason

; APPLICANT: Kowek, Allison

; TITLE OF INVENTION: THERAPEUTIC AND DIAGNOSTIC TOOLS FOR IMPAIRED GLUCOSE TOLERANCE CONDITIONS

; FILE REFERENCE: 00786/351005

; CURRENT APPLICATION NUMBER: US/09/844-353A

; CURRENT FILING DATE: 2001-04-27

; PRIOR APPLICATION NUMBER: US 08/857,076

; PRIOR FILING DATE: 1997-05-15

; NUMBER OF SEQ ID NOS: 114

; SOFTWARE: FastSEQ for Windows Version 4.0

; SEQ ID NO: 57

; LENGTH: 655

; TYPE: PRT

; ORGANISM: Homo sapiens

US-09-844-353A-57

Query Match 77.0%; Score 427.5; DB 9; Length 655;
Best Local Similarity 73.5%; Pred. No. 5.1e-42;

This Page Blank (uspto)

Copyright (c) 1993 - 2004: Compugen Ltd.

OM protein - protein search, using sw mode.

Run on: July 23, 2004, 10:22:15 ; search time 16 Seconds

Title: US-09-844-353A-54
Perfect score: 555
Sequence: 1 KKTTRRNAGNMSYAILIT.....SSWWVNPDAKPGMNPRTTR 103

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5
Searched: 283366 seqs, 96191526 residues

Total number of hits satisfying chosen parameters: 283366

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing First 45 summaries

Database : PIR 78:
1: pir1:
2: pir2:
3: pir3:
4: pir4:

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	54.9	98.9	510	2 T42255	HNF-3/forkhead protein homolog daf-16 - Caenorhabditis elegans
2	427.5	77.0	655	2 S40521	C;Species: Caenorhabditis elegans
3	372.5	67.1	530	2 T42234	C;Date: 03-Dec-1999 #sequence_revision 03-Dec-1999 #text_change 24-Aug-2001
4	372.5	67.1	622	2 T3204	C;Accession: T42255
5	322	58.0	235	2 T3203	R;Lin, K.; Dorman, J.B.; Rodan, A.; Kenyon, C.
6	209.5	37.7	230	2 T19437	Science 278: 1319-1322, 1997
7	195.5	35.2	270	2 T16880	A;Title: Daf-16, An HNF-3/forkhead family member that can function to double the life span of C. elegans
8	19.1	34.4	310	2 T33497	A;Reference number: 98028757; PMID:9360933
9	19.0	34.2	101	2 T33498	A;Status: preliminary; translated from GB/EMBL/DDJB
10	188.5	34.0	109	2 B54743	A;Molecule type: mRNA
11	188.5	34.0	387	2 A47446	A;Cross-references: EMBL:AF032112; NID:92623942; BIDN: AAC47803.1; PID:92623943
12	188.5	34.0	451	2 M55909	C;Genetics:
13	188.5	34.0	476	2 A54743	A;Gene: daf-16
14	188.5	34.0	480	2 J06072	C;Superfamily: unassigned fork head proteins; fork head DNA-binding domain homology
15	186	33.5	445	2 S23055	RESULT 1
16	186	33.5	445	2 S23056	T42255 HNF-3/forkhead protein homolog daf-16 - Caenorhabditis elegans
17	185.5	33.4	469	2 J37451	C;Species: Homo sapiens (man)
18	185	33.3	134	2 B46178	C;Date: 07-Oct-1994 #sequence_revision 03-Nov-1995 #text_change 07-May-1999
19	184.5	33.2	532	2 S2262	C;Accession: S40521
20	184	33.2	101	2 I60919	R;Galili, N.; Davis, R.J.; Fredericks, W.J.; Rauscher III, F.J.; Emanuel, A;Title: Fusion of a fork head domain gene to PAX3 in the solid tumour alveolar rhabdomy
21	183.5	33.1	586	2 JC6500	A;Reference number: S40521; PMID:9410095; PMID:8275086
22	183.5	33.1	663	2 J40493	A;Status: preliminary
23	182.5	32.9	101	2 I60922	A;Molecule type: mRNA
24	182.5	32.9	532	2 S51627	A;Residues: 1-655 <GAL>
25	182.5	32.9	106	2 S51626	C;Superfamily: unassigned fork head proteins; fork head DNA-binding domain homology
26	182.5	32.9	117	2 I49674	E160-248/Domain: fork head DNA-binding domain homology <FHD>
27	182.5	32.9	440	2 S71795	Query Match 98.9%; Score 549; DB 2; Pred. No. 3.e-49; Mismatches 0; Indels 0; Gaps 0;
28	182.5	32.9	465	2 G02738	Best Local Similarity 73.5%; Pred. No. 1.e-36; Mismatches 18; Indels 1; Gaps 1;
29	181.5	32.7	461	2 S34472	

Db	210	AAGWKNSIRHNLSLHSRPFMRQEGGAKSWWVNPDAKPGRNPRTR	257
		RESULT 5	
1	KKTTTRRNAGNMNSYAELETTAIMASPEKRLTIAQYEMMVQNVPYFRDKGDSNSAGWK	60	T37203 hypothetical protein R13H8.2 - Caenorhabditis elegans
151	KSSSRNAGNLSDAHLTKAIESSAERLTLTSQYBWNKSVPIFKDKGDSNSAGWK	210	C;Species: Caenorhabditis elegans C;Date: 03-Dec-1999 #sequence_revision 03-Dec-1999 #text_change 03-Dec-1999
61	NSIREHNLSLHSEWRNLQNEGAGKSSWWVNPDP-AKPGMNPFR 101		C;Accession: T37203 R;Jones, K.; Hinds, K.; Sutcher, C.; Cofman, M. R;Description: submitted to the EMBL Data Library, December 1997 A;Description: The sequence of C. elegans cosmid R13H8.
211	NSIREHNLSLHSEWRNLQNEGAGKSSWWVNPDP-AKPGMNPFR 252		A;Reference number: Z21633 A;Accession: T37203 A;Status: preliminary; translated from GB/EMBL/DDBJ
		RESULT 3	
1	KKTTTRRNAGNMNSYAELETTAIMASPEKRLTIAQYEMMVQNVPYFRDKGDSNSAGWK	60	T37234 fork head-related transcription factor homolog - Caenorhabditis elegans
	Species: Caenorhabditis elegans		C;Species: Caenorhabditis elegans C;Date: 03-Dec-1999 #sequence_revision 03-Dec-1999 #text_change 21-Jul-2000
	Accession: T42234		A;Cross-references: EMBL:AF039717; PIDN:AAE96742.1 A;Experimental source: strain Bristol N2 C;Genetics:
	(OGG, S.; Paradis, S.; Gottlieb, S.; Patterson, G.I.; Lee, L.; Tissenbaum, H.A.; Ryvkin Nature 369, 994-999, 1997 Title: The fork head transcription factor DAF-16 transduces insulin-like metabolic and Reference number: Z22208; MUID:98013175; PMID:9353126		A;Map position: 1 A;Introns: 82/2; 121/2; 168/1 A;Note: R13H8.2
		Query Match	Score 58.0%; DB 2; Length 235;
		Best Local Similarity	96.8%; Pred. No. 5.1e-26;
		Matches	61; Conservative 1; Mismatches 1; Indels 0; Gaps 0
		Qy	1 KKTTTRRNAGNMNSYAELETTAIMASPEKRLTIAQYEMMVQNVPYFRDKGDSNSAGWK 60
		Db	133 KKTTTRRNAGNMNSYAELETTAIMASPEKRLTIAQYEMMVQNVPYFRDKGDSNSAGWK 192
		Qy	61 NSI 63 Db 193 VSL 195
		RESULT 6	
1	KKTTTRRNAGNMNSYAELETTAIMASPEKRLTIAQYEMMVQNVPYFRDKGDSNS 55		T19437 hypothetical protein C25A1.2 - Caenorhabditis elegans
b	KKPTDQLAQKENPVGEEYSYTDIAKALESAPGRLKLNEIYQWSDNTIPYFRGESSPE 209		C;Species: Caenorhabditis elegans C;Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 15-Oct-1999
150	SAGWKNSIRHNLSLHSRMRQINEGAGKSWWVNPDAKPGRNPRTR 103		C;Accession: T19437 R;Mortimore, B. submitted to the EMBL Data Library, October 1996
56	SAGWKNSIRHNLSLHSRMRQINEGAGKSWWVNPDAKPGRNPRTR 257		A;Accession: T19437 A;Status: preliminary; translated from GB/EMBL/DDBJ
	Y	Score 67.1%; DB 2; Length 530;	A;Species: Caenorhabditis elegans A;Molecule type: DNA A;Residues: 1-230 <WIL>
	Y	Best Local Similarity 61.1%; Pred. No. 7.4e-31;	A;Cross-references: EMBL:Z81038; PIDN:CAB02761.1; GSPDB:GN00019; CESP:C25A1.2 A;Experimental source: clone C25A1 C;Genetics:
	Y	Matches 66; Conservative 15; Mismatches 22; Indels 5; Gaps 1;	A;Gene: CESP:C25A1.2 A;Map Position: 1 A;Introns: 35/2; 103/3; 146/3
	b	Query Match	Score 372.5%; DB 2; Length 235;
	b	Best Local Similarity	96.8%; Pred. No. 5.1e-26;
	b	Matches	61; Conservative 1; Mismatches 1; Indels 0; Gaps 0
		Qy	1 KKTTTRRNAGNMNSYAELETTAIMASPEKRLTIAQYEMMVQNVPYFRDKGDSNS 55
		Db	82 SYTGILAMATLSSPKQKMLAETWIMNEYPYFRSRG---AGWRNSIRHNLSDCF 134
		Qy	74 MRIQNEAGKSSWWVNPDAKPGRNPRTR 91
		Db	137 VRAGGRAANGKGHIVWAVHP 154
		RESULT 7	
1	KKTTTRRNAGNMNSYAELETTAIMASPEKRLTIAQYEMMVQNVPYFRDKGDSNS 55		T16880 hypothetical protein T14G12.4 - Caenorhabditis elegans
150	SAGWKNSIRHNLSLHSRMRQINEGAGKSWWVNPDAKPGRNPRTR 103		

C;Species: *Caenorhabditis elegans*
 C;Accession: 20-Sep-1999 #sequence_revision 20-Sep-1999 #text_change 21-Jan-2000
 R;Milcox, L.
 submitted to the EMBL Data Library, November 1995
 A;Reference number: Z16880
 A;Accession: T16880
 A;Status: preliminary; translated from GB/EMBL/DDBJ
 A;Molecule type: DNA
 A;Residues: 1-1270 <WIL>
 A;Cross-references: EMBL:U41268; NID:91086843; PID:91086847; PIDN:AAA82436.1; CESP:T14G1
 C;Genetics:
 A;Gene: CESP:T14G12.4
 A;Introns: 37/1; 72/3; 164/1
 F;93-185/Domain: fork head DNA-binding domain homology <FHD>
 Query Match 35.2% Score 195.5; DB 2; Length 270;
 Best Local Similarity 42.7%; Pred. No. 8e-13;
 Matches 41; Conservative 20; Mismatches 28; Indels 7; Gaps 2;
 Qy 1 KTTTTRRNAGNMSYAELITMASPEKRLLTAQVYEMVQNVPYFRDKGDNSSSAGWK 60
 Db 84 EKTPSPSPNDKEPPSYNALIIMMAIKDSPEKRLLTAQVYEMVQNVPYFRD-----NKGQWQ 138
 RESULT 10
 Qy 61 NSISTRNLSHSRMRION-EGAKGKSSWWVNPDAK 94
 Db 139 NSISTRNLSHSRMRIONCFVKVRPNFDDPGKGNYWMLDATAE 174

RESULT 8
 T33497 hypothetical protein F40H3.4 - *Caenorhabditis elegans*
 C;Species: *Caenorhabditis elegans*
 C;Date: 29-Oct-1999 #sequence_revision 29-Oct-1999 #text_change 29-Oct-1999
 R;Gattung, S.
 submitted to the EMBL Data Library, October 1998
 A;Description: The sequence of *C. elegans* cosmid F40H3.
 A;Accession: T33497
 A;Status: preliminary; translated from GB/EMBL/DDBJ
 A;Molecule type: DNA
 A;Residues: 1-310 <GAT>
 A;Experimental source: strain Bristol N2; clone F40H3
 C;Genetics:
 A;Map position: 2
 A;Gene: CESP:F40H3.4
 A;Introns: 37/3; 102/3; 126/1; 159/1; 186/3; 224/3; 258/1
 Query Match 34.4% Score 191; DB 2; Length 310;
 Best Local Similarity 42.0%; Pred. No. 2.8e-12;
 Matches 37; Conservative 20; Mismatches 23; Indels 8; Gaps 3;
 Qy 14 SYAELITMASPEKRLLTAQVYEMVQNVPYFRDKGDNSSSAGWKSTRHNLSLHSRF 73
 Db 60 SYSQTLRLATEDTPKKCTLAETYSPIAINQFYIRE---NRNSSLWNSTRHNLSLNKOF 115
 RESULT 11
 Qy 74 MRIQNEGAGKSSWWVNPDAK 100
 Db 116 SRIEKTDGDERGWWVCVDPPAK--KPR 140

RESULT 9
 A47450 HNF-3/fork head family transcription factor Qin - avian sarcoma virus 31
 C;Species: avian sarcoma virus 31, Asv31
 C;Accession: A47446
 R;Li, J.; Vogt, P.K.
 Proc. Natl. Acad. Sci. U.S.A. 90, 4490-4494, 1993
 A;Title: The retroviral oncogene qin belongs to the transcription factor family that inc
 A;Reference number: A47446
 A;Accession: A47446
 A;Status: preliminary; not compared with conceptual translation
 A;Molecule type: mRNA
 A;Residues: 1-387 <LII>

A;Cross-references: GB:Li10719; NID:9303962
A;Note: sequence extracted from NCBI backbone (NCBIP:133114)
A;Note: this translation is not annotated in GenBank entry ACSTRANS, release 113.0
C;Superfamily: unassigned fork head proteins; fork head DNA-binding domain homology
C;Keywords: transcription factor
F:151-242/Domain: fork head DNA-binding domain homology <FHD>

Query Match Score 188.5; DB 2; Length 387;
Best Local Similarity 43.9%; Pred. No. 6.3e-12;
Matches 20; Mismatches 19; Indels 7; Gaps 2;

Qy 14 SYAELITTAIMASPERKRLLTAAQYEMVQNVPYFRDKGDSNSSAGWKNSIRHNLSLHSRF 73
Db 155 SYNALIMMAIRQSPERKLTLNGIYEFLMKNPFPYRE----NKQGWQNSIRHNLSLNKCF 209

Qy 74 MRION - EGAGKSSWWVINPDA 93
Db 210 VKVPRHYDDPGKGNYWWMDPSS 231

RESULT 12
A55909 transforming protein c-qin - chicken
N;Alternate names: transcription factor CBF-1
C;Species: Gallus gallus (chicken)
C;Date: 23-Mar-1995 #sequence_revision 05-Apr-1995 #text_change 24-Sep-1999
C;Accession: A55909; S71794
R;Chang, H.W.; Li, J.; Kretzschmar, D.; Vogt, P.K.
Proc. Natl. Acad. Sci. U.S.A. 92, 447-451, 1995
A;Title: Avian cellular homolog of the qin Oncogene
A;Reference number: A55909; MUID:95132616; PMID:7831308
A;Accession: A55909
A;Molecule type: mRNA
A;Residues: 1-451 <CHA>
A;Cross-references: GB:L26814; NID:9642602; PIDN:AAA66954.1; PID:g642603
R;Yuasa, J.; Hirano, S.; Yamagata, M.; Noda, M.
Nature 382, 632-635, 1996
A;Title: Visual project map specified by topographic expression of transcription factor
A;Reference number: S71794; MUID:96338226; PMID:875714
A;Accession: S71794
A;Status: preliminary; nucleic acid sequence not shown; translation not shown
A;Molecule type: mRNA
A;Residues: 1-451 <YUA>
A;Cross-references: EMBL:U47275; NID:91546781; PIDN:AA08466.1; PID:g1546782
A;Note: the nucleotide sequence was submitted to the EMBL Data Library, January 1996
C;Superfamily: unassigned fork head proteins; fork head DNA-binding domain homology
C;Keywords: DNA binding; proto-oncogene
F:143-234/Domain: fork head DNA-binding domain homology <FHD>

Query Match Score 188.5; DB 2; Length 451;
Best Local Similarity 43.9%; Pred. No. 7.3e-12;
Matches 20; Mismatches 19; Indels 7; Gaps 2;

Qy 14 SYAELITTAIMASPERKRLLTAAQYEMVQNVPYFRDKGDSNSSAGWKNSIRHNLSLHSRF 73
Db 147 SYNALIMMAIRQSPERKLTLNGIYEFLMKNPFPYRE----NKQGWQNSIRHNLSLNKCF 201

Qy 74 MRION - EGAGKSSWWVINPDA 93
Db 202 VKVPRHYDDPGKGNYWWMDPSS 223

RESULT 13
A54743 transcription factor HFK1 - human
C;Species: Homo sapiens (man)
C;Accession: A54743
C;Cross-references: A54743
R;Murphy, D.B.; Wiese, S.; Burfeind, P.; Schmandt, D.; Mattei, M.G.; Schulz-Schaefter, W.
A;Title: Human brain factor 1, a new member of the Fork head gene family.
A;Reference number: A54743; MUID:7959731

A;Accession: A54743
A;Status: preliminary; nucleic acid sequence not shown
A;Molecule type: mRNA
A;Residues: 1-476
A;Cross-references: GB:X74142
C;Genetics:
A;Gene: GDB: FKH14 HFP-1; HFK1
A;Cross-references: GDB:431550
A;Map Position: 14q12-14q12
C;Superfamily: unassigned fork head proteins; fork head DNA-binding domain homology <FHD>
F:169-260/Domain: fork head DNA-binding domain homology <FHD>

Query Match Score 188.5; DB 2; Length 476;
Best Local Similarity 43.9%; Pred. No. 8.2e-12;
Matches 36; Conservative 20; Mismatches 19; Indels 7; Gaps 2;

Qy 14 SYAELITTAIMASPERKRLLTAAQYEMVQNVPYFRDKGDSNSSAGWKNSIRHNLSLHSRF 73
Db 173 SYNALIMMAIRQSPERKLTLNGIYEFLMKNPFPYRE----NKQGWQNSIRHNLSLNKCF 227

Qy 74 MRION - EGAGKSSWWVINPDA 93
Db 228 VKVPRHYDDPGKGNYWWMDPSS 249

RESULT 14
JH0672 brain factor 1 protein - rat
C;Species: Rattus norvegicus (Norway rat)
C;Accession: JH0672
R;Tao, W.; Lai, E.
Neuron 8, 957-966, 1992
A;Title: Telencephalon-restricted expression of BF-1, a new member of the HNF-3/fork head
A;Accession number: JH0672; MUID:92265309; PMID:1350202
A;Molecule type: mRNA
A;Residues: 1-480 <TAO>
A;Cross-references: GB:M87634; NID:9203134; PIDN:AAA40812.1; PID:g203135
A;Experimental source: brain
C;Superfamily: unassigned fork head proteins; fork head DNA-binding domain homology
C;Keywords: DNA binding; transcription factor
F:162-271/Region: DNA binding #status predicted
F:172-263/Domain: fork head DNA-binding domain homology <FHD>

Query Match Score 188.5; DB 2; Length 480;
Best Local Similarity 43.9%; Pred. No. 8.3e-12;
Matches 36; Conservative 20; Mismatches 19; Indels 7; Gaps 2;

Qy 14 SYAELITTAIMASPERKRLLTAAQYEMVQNVPYFRDKGDSNSSAGWKNSIRHNLSLHSRF 73
Db 176 SYNALIMMAIRQSPERKLTLNGIYEFLMKNPFPYRE----NKQGWQNSIRHNLSLNKCF 230

Qy 74 MRION - EGAGKSSWWVINPDA 93
Db 231 VKVPRHYDDPGKGNYWWMDPSS 252

RESULT 15
S23055 S1P2 protein - fruit fly (Drosophila melanogaster)
C;Species: Drosophila melanogaster
C;Date: 19-Feb-1994 #sequence_revision 01-Sep-1995 #text_change 24-Sep-1999
C;Accession: S23055
R;Grossniklaus, U.; Pearson, R. K.; Gehring, W.J.
Genes Dev. 6, 1030-1051, 1992
A;Title: The Drosophila slc08 Paired locus encodes two proteins involved in segmentation
A;Reference number: S23053; MUID:92275347; PMID:1317319
A;Accession: S23055
A;Molecule type: DNA
A;Residues: 1-445 <GRO>
A;Cross-references: EMBL:X66098; NID:98620; PIDN:CAA46892.1; PID:g6621

C;Genetics:
A;Gene: FlyBase:slp2
A;Cross-references: FlyBase:FBgn0004567
C;Superfamily: unassigned fork head proteins; fork head DNA-binding domain homology
C;Keywords: DNA binding; transcription regulation
F;180-271/Domain: fork head DNA-binding domain homology <FHD>

Query Match 33.5%; Score 186; DB 2; Length 445;
Best Local Similarity 41.8%; Pred. No. 1.4e-11; Mismatches 18; Indels 14; Gaps 3;
Matches 41; Conservative 18; Mismatches 23; Indels 14; Gaps 3;

Qy 14 SYAEPLITMASPEKRLLTIAQYEMMYCNVPYERDKGSNSAGWKNISIRHNLISLHSRF 73
Db 184 SYNALIMMAIROQSSEKRLLTFLNGIYBYIMTNHPYRD----NFGQGWQNSIRHNLISLNKCF 238

Qy 74 MRION-EGAGKSSWWVINDAK-----PGMNPRT 102
Db 239 VKVERHYDGPGRNYWMIDPSAEDVFIGGSTGKERRRT 276

Search completed: July 23, 2004, 10:25:54
Job time : 17 secs

This Page Blank (Upside)

?/1/?

Scoring table: BLASTM62							
Gapop 10.0 , Gapext 0.5							
Searched: 141681 seqs, 52070155 residues							
Total number of hits satisfying chosen parameters:							
141681							
Minimum DB seq length: 0							
Maximum DB seq length: 20000000000							
Post-processing: Minimum Match 0 %							
Maximum Match 10 %							
Listing first 45 summaries							
Database : Swissprot_42:*							
Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.							
SUMMARIES							
Result No.	Score		Query Match Length		DB ID	Description	
1	440.5		79.4		673	1 FX03_HUMAN	
2	427.5		77.0		655	1 FX01_HUMAN	
3	426.5		76.8		652	1 FX01_MOUSE	
4	413.5		74.5		505	1 FX04_MOUSE	
5	410.5		74.0		505	1 FX04_HUMAN	
6	222.5		40.1		622	1 YA41_HUMAN	
7	201.5		36.3		565	1 FXJ2_MOUSE	
8	190.5		36.3		574	1 FXJ2_HUMAN	
9	190.5		34.2		351	1 YK11_HUMAN	
10	188.5		34.2		743	1 YK18_SCHEPO	
11	188.5		34.0		387	1 QIN_AVIS3	
12	188.5		34.0		451	1 FXGB_CHICK	
13	188.5		34.0		477	1 FXGB_HUMAN	
14	188.5		34.0		480	1 FXGB RAT	
15	188.5		34.0		481	1 FXGB MOUSE	
16	187.5		33.8		365	1 PD2_DROME	
17	186.5		33.5		445	1 SLF2_DROME	
18	185.5		33.4		469	1 FXGA CHICK	
19	184.5		33.2		456	1 FXD1_MOUSE	
20	184.5		33.2		101	1 FXII_RAT	
21	183.5		33.1		663	1 SEP1_SCHEPO	
22	182.5		32.9		101	1 FXGA RAT	
23	182.5		32.9		440	1 FXGA CHICK	
24	182.5		32.9		465	1 FXD1_HUMAN	
25	182.5		32.9		497	1 FXD2_HUMAN	
26	182.5		32.9		553	1 FXC1_HUMAN	
27	182.5		32.9		763	1 FXC1_MOUSE	
28	182.5		32.8		Q08050	1 h forkhead	
29	181.5		32.7		494	1 FXC2_MOUSE	
30	181.5		32.5		501	1 FXC2_HUMAN	
31	180.5		32.5		759	1 FXM1 RAT	
32	180.5		32.5		760	1 FXM1 MOUSE	
33	180		32.4		490	1 FXN3_HUMAN	

ALIGNMENTS							
RESULT 1							
FX03_HUMAN							
ID	FX03_HUMAN		STANDARD;		PRT;	673 AA.	
AC	O43524; O15171; Q9BZ04;		PRT;			673 AA.	
DT	30-MAY-2000 (Rel. 3.9, Created)		PRT;			673 AA.	
DT	30-MAY-2000 (Rel. 3.9, Last sequence update)		PRT;			673 AA.	
DE	Forkhead box protein O3A (Forkhead in rhabdomyosarcoma-like 1) (AF6q21 protein).		PRT;			673 AA.	
DR	FOXO3A OR FXHLR1.		PRT;			673 AA.	
OS	Homo sapiens (Human).		PRT;			673 AA.	
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;		PRT;			673 AA.	
OX	NCBI_TaxID=9606;		PRT;			673 AA.	
RN	[1]		SEQUENCE FROM N.A.			673 AA.	
RP	SEQUENCE FROM N.A.		PRT;			673 AA.	
RX	MEDLINE=98140118; PubMed=9479491;		PRT;			673 AA.	
RA	Anderson M.J., Viars C.S., Czekay S., Cavenee W.K., Arden K.C.,		PRT;			673 AA.	
RT	"Cloning and characterization of three human forkhead genes that comprise an FXHR-like gene subfamily."		PRT;			673 AA.	
RT	Genomics 47:187-199 (1998).		PRT;			673 AA.	
RL	[2]		SEQUENCE FROM N.A.			673 AA.	
RN	[2]		SEQUENCE FROM N.A.			673 AA.	
RP	SEQUENCE FROM N.A.		PRT;			673 AA.	
RA	Cobley V.		PRT;			673 AA.	
RL	Submitted (DEC-2000) to the EMBL/GenBank/DDBJ databases.		PRT;			673 AA.	
RN	[3]		SEQUENCE FROM N.A.			673 AA.	
RP	SEQUENCE FROM N.A.		PRT;			673 AA.	
RC	TISSUE: Muscle;		PRT;			673 AA.	
RX	MEDLINE=22388257; PubMed=12477932;		PRT;			673 AA.	
RA	Straubberg R.L., Feingold E.A., Grouse L.H., Derge J.G., Derge J.G., Schuler G.D., Klausner R.D., Collins F.S., Wagner L., Sheinman C.M., Schuler G.D., Altschuler S.F., Zeeberg B., Buetow K.H., Schaefer C.P., Bhat N.K., Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F., Diatchenko L., Marusina K., Farmer A.A., Rubin M.F., Casavant T.L., Scheetz T.E., Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E., Brownstein M.J., Ustün T.B., Toshiyuki S., Carninci P., Prange C., Blakesley R.W., Logue N.A., Abramson R.D., Millahy S.J., Bosak S.A., McEwan P.J., Petersen K.J., Malek J.A., Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W., Villain D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A., Fahey J., Hellion E., Kettman M., Madan A., Rodriguez S., Sanchez A., Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G., Rodriguez A.C., Grimwood J.W., Green E.D., Gunaratne P.H., Butterfield Y.S.N., Krzywinski M.I., Skalska U., Smalius D.E., Schmid A., Schein J., Jones S.J.M., Marra M.A., Prokunina-Olsson M., Generation and initial analysis of more than 15,000 full-length human and mouse cDNA sequences.", Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903 (2002).		PRT;			673 AA.	
RN	[4]		SEQUENCE OF 1-383 FROM N.A.			673 AA.	
RP	MEDLINE=9800138; PubMed=9345057;		PRT;			673 AA.	
RA	Hallion J., Le Coniat M., Jonveaux P., Berger R., Bernard O.A., "Af6q21, a novel partner of the MLL gene in t(6;11) (q21;q23), defines a Forkhead transcriptional factor subfamily."		PRT;			673 AA.	
RT	Blood 90:3714-3719(1997).		PRT;			673 AA.	
RL	Blood 90:3714-3719(1997).		PRT;			673 AA.	

[5] PHOSPHORYLATION ON SER-315 BY SGK.
 RP MEDLINE=1154281;
 RX J. Tran H., Hu L.S., Hemmings B.A., Greenberg M.E.;
 RA Brunet A., Park J.Y., Tran H., Hu L.S., Hemmings B.A., Greenberg M.E.;
 RT 'protein kinase SGK mediates survival signals by phosphorylating the
 forkhead transcription factor FOXR1L (FOXO3a).';
 RU Mol. Cell. Biol. 21:952-965 (2001).
 CC -!- FUNCTION: Probable transcription factor that may trigger apoptosis
 CC by inducing the expression of genes that are critical for cell
 CC death.
 CC -!- SUBCELLULAR LOCATION: Nuclear.
 CC -!- TISSUE SPECIFICITY: ubiquitous.
 CC -!- PTM: Phosphorylated by AKT1 and by SGK/SGK1.
 CC -!- DISEASE: Involved in secondary acute leukemias by a chromosomal
 translocation t(6;11) (q1;q2), that involves FOXO3A and MLL/hox.
 CC -!- SIMILARITY: Contains 1 fork-head domain.
 CC -!- CAUTION: Ref.2 sequence differs from that shown due to erroneous
 Gene model prediction.
 CC -!- DATABASE: NAMEAtlas Genet. Cytogenet. Oncol. Haematol.;
 WWW="http://www.infobiogen.fr/services/chromcancer/Genes/AF6q21.1D125.html".
 CC -!- SWISS-PROT entry is copyright. It is produced through a collaboration
 between the Swiss Institute of Bioinformatics and the EMBL outstation -
 European Bioinformatics Institute. There are no restrictions on its
 use by non-profit institutions as long as its content is in no way
 modified and this statement is not removed. Usage by and for commercial
 entities requires a license agreement (See http://www.isb-sib.ch/announce/
 or send an email to license@isb-sib.ch).
 CC -!- EMBL; AF022886; AAC39592.1; ALT_SEQ.
 DR EMBL; AL265509; CAC26821.1; ALT_SEQ.
 DR EMBL; BC02227; AAH20227.1; -.
 DR EMBL; BC01224; AAH21224.1; -.
 DR EMBL; AJ001589; CAA04860.1; -.
 DR EMBL; AJ001590; CAA04861.1; -.
 DR HSSP; Q6345; ZHFH.
 DR TRANSFAC; T02938; -.
 DR Gene; HGNC; 3821; FOXO3A.
 DR MIM; 602681; -.
 DR GO; GO:0005737; C:cytoplasm; TAS.
 DR GO; GO:0005634; C:nucleus; TAS.
 DR GO; GO:0006917; P:induction of apoptosis; TAS.
 DR SMART; SM0033; FH_1.
 DR PROSITE; PS00657; FORK_HEAD_1; FALSE_NEG.
 DR PROSITE; PS00658; FORK_HEAD_2; 1.
 DR PROSITE; PS00309; FORK_HEAD_3; 1.
 KW Transcription regulation; DNA-binding; Nuclear protein; Apoptosis.
 KW Chromosomal translocation; Proto-oncogene; Phosphorylation.
 KW DNA BIND
 MOD_RES 157 251 FORK_HEAD_4.
 MOD_RES 253 253 PHOSPHORYLATION (BY AKT1) (BY
 SIMILARITY).
 FT MOD_RES 315 315 PHOSPHORYLATION (BY SGK).
 FT CONFLICT 156 163 ANGNIKYA -> WKGPKYS (IN REF. 4).
 FT CONFLICT 238 246 PDGFKIGKA -> LMGGERKT (IN REF. 4).
 FT CONFLICT 253 253 S -> T (IN REF. 4).
 FT CONFLICT 271 271 MISSING (IN REF. 4).
 FT CONFLICT 292 330 PGSTTSRSDELLAWDIDFRSRISNASTVSGRLSPIMAS
 -> AWQPHVNAATMSWWRGRTSYHAPLTTPAQVAARPSW
 FT CONFLICT 315 315 QV (IN REF. 4).
 FT CONFLICT 345 361 PMLYSSASLSPVSKP -> AHALQAVSQPTVFSKQA
 FT CONFLICT 367 367 (IN REF. 4).
 FT CONFLICT 371 371 P -> R (IN REF. 4).
 FT CONFLICT 382 383 D -> E (IN REF. 4).
 FT CONFLICT 383 383 LT -> AD (IN REF. 4).
 SQ SEQUENCE 673 AA; 71276 MW; E5B4E830665A9982 CRC64;

	Matches	79; Conservative	13; Mismatches	9; Indels	1; Gaps	1;
QY	1	KCTTTRNAWNMGNSYALITTAIMASPERKLTLAQYEMWVNPYFRDKGDSNSAGWK 60				
Db	148	RKCSRRAWNAGWNLYSALITRAIESSEPMRNPDRKTLTQIYWMVRCVPYFKDKGDSNSAGWK 207				
QY	61	NSTRHNLSLHSRMRMIONEGACKSMMWVNPDA-KPGMNPRR 101				
Db	208	NSTRHNLSLHSRMRVNEGTGKSSWWINPDGKGSKAPRR 249				
RESULT 2						
EXO1_HUMAN STANDARD; PRT; 655 AA.						
AC	012778; O43523;					
DT	01-NOV-1997 (Rel. 35, Created)					
DT	01-NOV-1997 (Rel. 35, Last sequence update)					
DT	10-OCT-2003 (Rel. 42, Last annotation update)					
DB	Forkhead box protein O1A (Forkhead in rhabdomyosarcoma).					
GN	FOXA1 OR FKHR.					
OS	<i>Homo sapiens</i> (Human).					
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;					
MA	Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.					
NCBI_TaxID	9606;					
RN						
RP	SEQUENCE FROM N.A.					
RX	MEDLINE=4100975; PubMed=8275086;					
RA	Galli N., Davis R.J., Fredericks W.J., Mukhopadhyay S.,					
RA	Rauscher F.J. III, Emanuel B.S., Rovera G., Barr F.G.,					
RT	"Fusion of a Fork head domain gene to PAX3 in the solid tumour alveolar rhabdomyosarcoma."					
RL	Nat. Genet. 5:230-235 (1993).					
RN	[2]					
RP	SEQUENCE FROM N.A.					
RC	TISSUE=Lymph;					
RX	MEDLINE=98140118; PubMed=9479491;					
RA	Anderson M.J., Viars C.S., Czekay S., Cavenee W.K., Arden K.C.,					
RT	"Cloning and characterization of three human forkhead genes that comprise an FKHR-like gene subfamily."					
RL	Genomics 47:187-199 (1998).					
RN	[3]					
RP	SEQUENCE FROM N.A.					
RC	TISSUE=Lymph;					
RX	MEDLINE=22388257; PubMed=12477932;					
RA	Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G., Schuler G.D., Klaushier R.D., Collins F.S., Wagner L., Shmueli C.M., Schulz C.P., Bhat N.K., Altschul S.F., Zeeberg B., Buston K.H., Schaefer C.F., Hsieh F., Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hong L., Diatchenko L., Marusina K., Farmer A.A., Rubin M.F., Bonaldo M.F., Scheetz T.E., Stapleton M., Soares M.B., Savavant T.L., Carninci P., Prange C., Brownstein M.J., Usdin T.B., Toshiyuki S., Carninci P., Mullahy S.J., Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Gunaratne P.H., Bosak S.A., McEvlan P.J., McKernan K.J., Malek J.A., Hale S.S., Garcia A.M., Gay L.J., Hulyk S.W., Richards S., Worley K.C., Villalon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A., Fahy J., Heitton B., Kettman M., Madan A., Rodriguez S., Sanchez A., Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G., Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C., Rodríguez A.C., Grimwood J., Schmutz J., Myers R.M., Rutter M., Butterfield Y.S.N., Krzywinski M.I., Skalska U., Smilis D.E., Schnarch A., Schein J.E., Jones S.J.M., Marra M.A., Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903 (2002).					
CC	-!- FUNCTION: Probable transcription factor.					
CC	-!- SUBCELLULAR LOCATION: Nuclear (By similarity).					
CC	-!- TISSUE SPECIFICITY: Ubiquitous.					
CC	-!- PIM: Phosphorylated by AKT1. Insulin-induced (By similarity).					
CC	-!- DISEASE: Involved in t(2;13) (q35;q14) and t(1;13) (p36;q14).					
CC	-!- chromosomal translocations in alveolar rhabdomyosarcoma-2 (RMS2)					
CC	that involves FOXA1 and PAX3 or PAX7. The resulting protein is a transcriptional activator.					
CC	-!- SIMILARITY: Contains 1 fork-head domain.					

This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to licensee@isb-sib.ch).

DR EMBL; U03310; AAA3629; 1; -.

DR EMBL; AF032885; AAC39591; 1; -.

DR EMBL; BC021981; AHR21981; 1; -.

DR PIR; S40521; S40521; -.

DR HSSP; Q03245; 2HEH; -.

DR GeneID; HGNC:3819; FOXO1A.

DR MIM; 130533; -.

DR GO; GO:003700; F transcription factor activity; TAS.

DR GO; GO:0006946; P anti-apoptosis; TAS.

DR GO; GO:0006357; P regulation of transcription from RNA polymerase II promoter; IPR001786; TF_Fork_head; 1.

DR InterPro; PR00250; Fork_head; 1.

DR Pram; PR00053; FORKHEAD.

DR SMART; SM00339; FH; 1.

DR PROSITE; PS000425; TF_Fork_head; 1.

DR PROSITE; PS00657; FORK_HEAD_1; FALSE_NEG.

DR PROSITE; PS00658; FORK_HEAD_2; 1.

DR PROSITE; PS50039; FORK_HEAD_3; 1.

KW Phosphorylation; Chromosomal translocation; Proto-oncogene.

FT DOMAIN_91 102 POLY-ALA.

FT DOMAIN_120 130 POLY-PRO.

FT DOMAIN_152 155 POLY-SER.

FT DOMAIN_159 235 FORK-HEAD.

FT MOD_RES 256 256 PHOSPHORYLATION (BY PKB/AKT1) (BY CONFLICT).

FT CONFFLICT 131 131 V > L (IN REF. 2).

SEQUENCE 655 AA; 69647 MW; 6DEFC994E740399 CRC64;

Query Match Score 77.0%; Best Local Similarity 73.5%; Matches 75; Conservative 18; Mismatches 8; Indels 1; Gaps 1;

Db 1 KTTTTRNAAGNMYSALITTAIAASPEKRLLTAAQYEMMYVNPYFRDKGDSNSAGWK 60

Db 151 KSSSSRNNAGNLYSADLITRAESEAKRLLSQIYEMMYVNPYFRDKGDSNSAGWK 210

Qy 61 NSIRHNLSHSRPMRILQNEGAGKSSMWVIND-AKGGMNPR 101

Db 211 NSIRHNLSHSKFIYQNEGTGKSSMWLNPEGGSGKSRR 252

RESULT 3

FXO1_MOUSE ID FXO1_MOUSE STANDARD; PRT; 652 AA.

AC Q9WTF3; OS Mus musculus (Mouse).

DT 30-MAY-2000 (Rel. 39, Created)

DT 30-MAY-2000 (Rel. 39, Last sequence update)

DT 28-FEB-2003 (Rel. 41, Last annotation update)

DE Forkhead box protein O1A (Forkhead in rhabdomyosarcoma).

GN FOXO1A OR FOXO1 OR FKHR.

OS Mus musculus (Mouse).

OC Mammalia; Chordata; Craniata; Vertebrata; Buteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.

OX NCBI_TaxID=10090;

RN SEQUENCE FROM N.A.

RX MEDLINE=927B156; PubMed=10347145;

RA Nakae J.; Park B.C.; Accili D.; RT "Insulin stimulates phosphorylation of the forkhead transcription factor FOXR on serine 253 through a Wortmannin-sensitive pathway." RL J. Biol. Chem. 274:15982-15985 (1999).

CC -1- FUNCTION: Probable transcription factor.

CC -1- SUBCELLULAR LOCATION: Nuclear (Potential).

CC -1- SIMILARITY: Contains 1 fork-head domain.

-1- PTM: PHOSPHORYLATED BY AKT; INSULIN-INDUCED.
-1- SIMILARITY: Contains 1 fork-head domain.

This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to licensee@isb-sib.ch).

DR EMBL; AF126056; AAD40636; 1; -.

DR HSSP; Q63245; 2HFH.

DR MGDB; MGII:1890077; Foxo1.

DR InterPro; IPR001766; TF_Fork_head.

DR Pfam; PF00250; Fork_head; 1.

DR PRINTS; PR00053; FORKHEAD.

DR PRODOM; PD000425; TF_Fork_head; 1.

DR SMART; SM00339; FH; 1.

DR PROSITE; PS00657; FORK_HEAD_1; FALSE_NEG.

DR PROSITE; PS00658; FORK_HEAD_2; 1.

DR PROSITE; PS50039; FORK_HEAD_3; 1.

KW Transcription regulation; DNA-binding; Nuclear protein; Phosphorylation.

FT DOMAIN 89 96 POLY-ALA.

FT DOMAIN 135 139 POLY-ALA.

FT DNA_BIND 156 232 FORK-HEAD.

FT MOD_RES 253 253 PHOSPHORYLATION (BY PKB/AKT1).

SQ SEQUENCE 652 AA; 69502 MW; 3FE4C322AA85205F CRC64;

Query Match Score 76.8%; Best Local Similarity 73.5%; Matches 75; Conservative 18; Mismatches 8; Indels 1; Gaps 1;

Qy 1 KTTTTRNAAGNMYSALITTAIAASPEKRLLTAAQYEMMYVNPYFRDKGDSNSAGWK 60

Db 148 KTTTTRNAAGNMYSALITRAESEAKRLLSQIYEMMYVNPYFRDKGDSNSAGWK 207

Query Match Score 427.5%; Best Local Similarity 73.5%; Matches 75; Conservative 18; Mismatches 8; Indels 1; Gaps 1;

Db 61 NSIRHNLSHSRPMRILQNEGAGKSSMWVIND-AKGGMNPR 101

Db 208 NSIRHNLSHSKFIYQNEGTGKSSMWLNPEGGSGKSRR 249

RESULT 4

FOX4_MOUSE ID FOX4_MOUSE STANDARD; PRT; 505 AA.

AC Q9WTF3; OS Mus musculus (Mouse).

DT 16-OCT-2001 (Rel. 40, Created)

DT 16-OCT-2001 (Rel. 40, Last sequence update)

DT 28-FEB-2003 (Rel. 41, Last annotation update)

DE Putative fork head domain transcription factor AFX1 (Foxh) (Foxzhead box protein 04).

GN ML17 OR AFX1 OR FOX4.

RP SEQUENCE FROM N.A.

RA Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.

OC NCBITaxID=10090;

RN [1]

RT "Identification and characterization of murine members of the FKHR subclass of winged-helix transcription factors." RL submitted (DRC-1998) to the EMBL/GenBank/DBJ databases.

RN [2]

RP SEQUENCE FROM N.A.

RA Furukama T.; Nakazawa T.; Mori N.;

RT "Mouse AF4, a forkhead type transcription factor." RL Submitted (SPB-1999) to the EMBL/GenBank/DBJ databases.

CC -1- FUNCTION: Plays a role in the insulin signaling pathway (By similarity).

CC -1- SUBCELLULAR LOCATION: Nuclear (Potential).

CC -1- SIMILARITY: Contains 1 fork-head domain.

This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and thus statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to license@isb-sib.ch).

CC MEDLINE=21864263; PubMed=11779849;
 CC RA "An mRNA splice variant of the AFX gene with altered transcriptional activity.";
 CC RT J. Biol. Chem. 277:8068-8075 (2002).
 CC [4]
 CC RN CHROMOSOMAL TRANSLOCATION.
 CC RP TLSUB-Bone marrow; PubMed=7529552;
 CC RX MEDLINE=95118921; PubMed=7529552;
 DR Parry P.; Wei Y.; Evans G.;
 RA "Cloning and characterization of the t(X;11) breakpoint from a leukemic cell line identify a new member of the forkhead gene family.";
 RT J. Clin. Oncol. 17:103-109 (1999).
 DR Genes Chromosomes Cancer 11:79-84 (1994).
 DR -1- FUNCTION: Plays a role in the insulin signaling pathway.
 DR -1- SUBCELLULAR LOCATION: Nuclear (Potential).
 DR -1- ALTERNATIVE PRODUCTS: Named isoforms=2;
 DR Event=Alternative splicing; Named isoforms=2;
 DR Name=1; Synonyms=FOXO4a;
 DR IsoID=P91177-1; Sequence=Displayed;
 DR Name=Zeta; Synonyms=AFZeta, FOXOb;
 DR IsoID=998177-2; Sequence=Y8P 001552;
 DR -1- TISSUE SPECIFICITY: Heart, brain, placenta, lung, liver, skeletal muscle, kidney and pancreas. Isoform zeta is most abundant in the liver, kidney, and pancreas.
 DR -1- DISEASE: Involved in acute leukemias by a chromosomal translocation t(X;11) (q13;q23) that involves MLL7 and MLL/HRX.
 DR The result is a rogue activator protein.
 DR -1- SIMILARITY: Contains 1 fork-head domain.
 DR -1- DATABASE: NAME=Atlas Genet. Cytogenet. Oncol. Haematol.;
 DR WWW=<http://www.infobiogen.fr/services/chromcancer/Genes/AFX1ID57.html>.
 DR ---

This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and thus statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to license@isb-sib.ch).

CC MEDLINE=9811552; PubMed=9811552;
 CC -1- TISSUE SPECIFICITY: Heart, brain, placenta, lung, liver, skeletal muscle, kidney and pancreas. Isoform zeta is most abundant in the liver, kidney, and pancreas.
 CC -1- DISEASE: Involved in acute leukemias by a chromosomal translocation t(X;11) (q13;q23) that involves MLL7 and MLL/HRX.
 CC The result is a rogue activator protein.
 CC -1- SIMILARITY: Contains 1 fork-head domain.
 CC -1- DATABASE: NAME=Atlas Genet. Cytogenet. Oncol. Haematol.;
 CC WWW=<http://www.infobiogen.fr/services/chromcancer/Genes/AFX1ID57.html>.
 CC ---

Query Match 74.5%; Score 413 5; DB 1; Length 505;
 Best Local Similarity 72.5%; Pred. No. 2.4e-38; Indels 1; Gaps 1;
 Matches 74; Conservative 14; Mismatches 13; Indels 1;

QY 1 KRTTPRNANMCSAELITTAIMASPEKRLTLAQYENMVQVYFDRKGDSNNSAAGWK 60
 :|:||:||:||:||:||:||:||:||:||:||:||:||:||:||:||:||:||:||:||:||:||:||:||:
 QY 92 RKGGSRRANQNOSTAELISQAESAPEKRLTLAQYENMVRTVYFDRKGDSNNSAAGWK 151
 KW Transcription regulation; DNA-binding; Nuclear protein.

FT DNA_BIND 100 188 FORK_HEAD.

FT SEQUENCE 505 AA; ABB99B54807CTCE5 CRC64;

FT PRODOM PD000425; FORK_HEAD; 1.

FT PROSITE PS00339; PH; 1.

FT PROSITE PS00657; FORK_HEAD; 1; FALSE_NEG.

FT PROSITE PS00658; FORK_HEAD; 2; 1.

FT PROSITE PS50039; FORK_HEAD; 3; 1.

FT PROSITE PS50039; FORK_HEAD; 4; 1.

FT PROSITE PS50039; FORK_HEAD; 5; 1.

FT PROSITE PS50039; FORK_HEAD; 6; 1.

FT PROSITE PS50039; FORK_HEAD; 7; 1.

FT PROSITE PS50039; FORK_HEAD; 8; 1.

FT PROSITE PS50039; FORK_HEAD; 9; 1.

FT PROSITE PS50039; FORK_HEAD; 10; 1.

FT PROSITE PS50039; FORK_HEAD; 11; 1.

FT PROSITE PS50039; FORK_HEAD; 12; 1.

FT PROSITE PS50039; FORK_HEAD; 13; 1.

FT PROSITE PS50039; FORK_HEAD; 14; 1.

FT PROSITE PS50039; FORK_HEAD; 15; 1.

FT PROSITE PS50039; FORK_HEAD; 16; 1.

FT PROSITE PS50039; FORK_HEAD; 17; 1.

FT PROSITE PS50039; FORK_HEAD; 18; 1.

FT PROSITE PS50039; FORK_HEAD; 19; 1.

FT PROSITE PS50039; FORK_HEAD; 20; 1.

FT PROSITE PS50039; FORK_HEAD; 21; 1.

FT PROSITE PS50039; FORK_HEAD; 22; 1.

FT PROSITE PS50039; FORK_HEAD; 23; 1.

FT PROSITE PS50039; FORK_HEAD; 24; 1.

FT PROSITE PS50039; FORK_HEAD; 25; 1.

FT PROSITE PS50039; FORK_HEAD; 26; 1.

FT PROSITE PS50039; FORK_HEAD; 27; 1.

FT PROSITE PS50039; FORK_HEAD; 28; 1.

FT PROSITE PS50039; FORK_HEAD; 29; 1.

FT PROSITE PS50039; FORK_HEAD; 30; 1.

FT PROSITE PS50039; FORK_HEAD; 31; 1.

FT PROSITE PS50039; FORK_HEAD; 32; 1.

FT PROSITE PS50039; FORK_HEAD; 33; 1.

FT PROSITE PS50039; FORK_HEAD; 34; 1.

FT PROSITE PS50039; FORK_HEAD; 35; 1.

FT PROSITE PS50039; FORK_HEAD; 36; 1.

FT PROSITE PS50039; FORK_HEAD; 37; 1.

FT PROSITE PS50039; FORK_HEAD; 38; 1.

FT PROSITE PS50039; FORK_HEAD; 39; 1.

FT PROSITE PS50039; FORK_HEAD; 40; 1.

FT PROSITE PS50039; FORK_HEAD; 41; 1.

FT PROSITE PS50039; FORK_HEAD; 42; 1.

FT PROSITE PS50039; FORK_HEAD; 43; 1.

FT PROSITE PS50039; FORK_HEAD; 44; 1.

FT PROSITE PS50039; FORK_HEAD; 45; 1.

FT PROSITE PS50039; FORK_HEAD; 46; 1.

FT PROSITE PS50039; FORK_HEAD; 47; 1.

FT PROSITE PS50039; FORK_HEAD; 48; 1.

FT PROSITE PS50039; FORK_HEAD; 49; 1.

FT PROSITE PS50039; FORK_HEAD; 50; 1.

FT PROSITE PS50039; FORK_HEAD; 51; 1.

FT PROSITE PS50039; FORK_HEAD; 52; 1.

FT PROSITE PS50039; FORK_HEAD; 53; 1.

FT PROSITE PS50039; FORK_HEAD; 54; 1.

FT PROSITE PS50039; FORK_HEAD; 55; 1.

FT PROSITE PS50039; FORK_HEAD; 56; 1.

FT PROSITE PS50039; FORK_HEAD; 57; 1.

FT PROSITE PS50039; FORK_HEAD; 58; 1.

FT PROSITE PS50039; FORK_HEAD; 59; 1.

FT PROSITE PS50039; FORK_HEAD; 60; 1.

FT PROSITE PS50039; FORK_HEAD; 61; 1.

FT PROSITE PS50039; FORK_HEAD; 62; 1.

FT PROSITE PS50039; FORK_HEAD; 63; 1.

FT PROSITE PS50039; FORK_HEAD; 64; 1.

FT PROSITE PS50039; FORK_HEAD; 65; 1.

FT PROSITE PS50039; FORK_HEAD; 66; 1.

FT PROSITE PS50039; FORK_HEAD; 67; 1.

FT PROSITE PS50039; FORK_HEAD; 68; 1.

FT PROSITE PS50039; FORK_HEAD; 69; 1.

FT PROSITE PS50039; FORK_HEAD; 70; 1.

FT PROSITE PS50039; FORK_HEAD; 71; 1.

FT PROSITE PS50039; FORK_HEAD; 72; 1.

FT PROSITE PS50039; FORK_HEAD; 73; 1.

FT PROSITE PS50039; FORK_HEAD; 74; 1.

FT PROSITE PS50039; FORK_HEAD; 75; 1.

FT PROSITE PS50039; FORK_HEAD; 76; 1.

FT PROSITE PS50039; FORK_HEAD; 77; 1.

FT PROSITE PS50039; FORK_HEAD; 78; 1.

FT PROSITE PS50039; FORK_HEAD; 79; 1.

FT PROSITE PS50039; FORK_HEAD; 80; 1.

FT PROSITE PS50039; FORK_HEAD; 81; 1.

FT PROSITE PS50039; FORK_HEAD; 82; 1.

FT PROSITE PS50039; FORK_HEAD; 83; 1.

FT PROSITE PS50039; FORK_HEAD; 84; 1.

FT PROSITE PS50039; FORK_HEAD; 85; 1.

FT PROSITE PS50039; FORK_HEAD; 86; 1.

FT PROSITE PS50039; FORK_HEAD; 87; 1.

FT PROSITE PS50039; FORK_HEAD; 88; 1.

FT PROSITE PS50039; FORK_HEAD; 89; 1.

FT PROSITE PS50039; FORK_HEAD; 90; 1.

FT PROSITE PS50039; FORK_HEAD; 91; 1.

FT PROSITE PS50039; FORK_HEAD; 92; 1.

FT PROSITE PS50039; FORK_HEAD; 93; 1.

FT PROSITE PS50039; FORK_HEAD; 94; 1.

FT PROSITE PS50039; FORK_HEAD; 95; 1.

FT PROSITE PS50039; FORK_HEAD; 96; 1.

FT PROSITE PS50039; FORK_HEAD; 97; 1.

FT PROSITE PS50039; FORK_HEAD; 98; 1.

FT PROSITE PS50039; FORK_HEAD; 99; 1.

FT PROSITE PS50039; FORK_HEAD; 100; 1.

FT PROSITE PS50039; FORK_HEAD; 101; 1.

FT PROSITE PS50039; FORK_HEAD; 102; 1.

FT PROSITE PS50039; FORK_HEAD; 103; 1.

FT PROSITE PS50039; FORK_HEAD; 104; 1.

FT PROSITE PS50039; FORK_HEAD; 105; 1.

FT PROSITE PS50039; FORK_HEAD; 106; 1.

FT PROSITE PS50039; FORK_HEAD; 107; 1.

FT PROSITE PS50039; FORK_HEAD; 108; 1.

FT PROSITE PS50039; FORK_HEAD; 109; 1.

FT PROSITE PS50039; FORK_HEAD; 110; 1.

FT PROSITE PS50039; FORK_HEAD; 111; 1.

FT PROSITE PS50039; FORK_HEAD; 112; 1.

FT PROSITE PS50039; FORK_HEAD; 113; 1.

FT PROSITE PS50039; FORK_HEAD; 114; 1.

FT PROSITE PS50039; FORK_HEAD; 115; 1.

FT PROSITE PS50039; FORK_HEAD; 116; 1.

FT PROSITE PS50039; FORK_HEAD; 117; 1.

FT PROSITE PS50039; FORK_HEAD; 118; 1.

FT PROSITE PS50039; FORK_HEAD; 119; 1.

FT PROSITE PS50039; FORK_HEAD; 120; 1.

FT PROSITE PS50039; FORK_HEAD; 121; 1.

FT PROSITE PS50039; FORK_HEAD; 122; 1.

FT PROSITE PS50039; FORK_HEAD; 123; 1.

FT PROSITE PS50039; FORK_HEAD; 124; 1.

FT PROSITE PS50039; FORK_HEAD; 125; 1.

FT PROSITE PS50039; FORK_HEAD; 126; 1.

FT PROSITE PS50039; FORK_HEAD; 127; 1.

FT PROSITE PS50039; FORK_HEAD; 128; 1.

FT PROSITE PS50039; FORK_HEAD; 129; 1.

FT PROSITE PS50039; FORK_HEAD; 130; 1.

FT PROSITE PS50039; FORK_HEAD; 131; 1.

FT PROSITE PS50039; FORK_HEAD; 132; 1.

FT PROSITE PS50039; FORK_HEAD; 133; 1.

FT PROSITE PS50039; FORK_HEAD; 134; 1.

FT PROSITE PS50039; FORK_HEAD; 135; 1.

FT PROSITE PS50039; FORK_HEAD; 136; 1.

FT PROSITE PS50039; FORK_HEAD; 137; 1.

FT PROSITE PS50039; FORK_HEAD; 138; 1.

FT PROSITE PS50039; FORK_HEAD; 139; 1.

FT PROSITE PS50039; FORK_HEAD; 140; 1.

FT PROSITE PS50039; FORK_HEAD; 141; 1.

FT PROSITE PS50039; FORK_HEAD; 142; 1.

FT PROSITE PS50039; FORK_HEAD; 143; 1.

FT PROSITE PS50039; FORK_HEAD; 144; 1.

FT PROSITE PS50039; FORK_HEAD; 145; 1.

FT PROSITE PS50039; FORK_HEAD; 146; 1.

FT PROSITE PS50039; FORK_HEAD; 147; 1.

FT PROSITE PS50039; FORK_HEAD; 148; 1.

FT PROSITE PS50039; FORK_HEAD; 149; 1.

FT PROSITE PS50039; FORK_HEAD; 150; 1.

FT PROSITE PS50039; FORK_HEAD; 151; 1.

FT PROSITE PS50039; FORK_HEAD; 152; 1.

FT PROSITE PS50039; FORK_HEAD; 153; 1.

FT PROSITE PS50039; FORK_HEAD; 154; 1.

FT PROSITE PS50039; FORK_HEAD; 155; 1.

FT PROSITE PS50039; FORK_HEAD; 156; 1.

FT PROSITE PS50039; FORK_HEAD; 157; 1.

FT PROSITE PS50039; FORK_HEAD; 158; 1.

FT PROSITE PS50039; FORK_HEAD; 159; 1.

FT PROSITE PS50039; FORK_HEAD; 160; 1.

FT PROSITE PS50039; FORK_HEAD; 161; 1.

FT PROSITE PS50039; FORK_HEAD; 162; 1.

FT PROSITE PS50039; FORK_HEAD; 163; 1.

FT PROSITE PS50039; FORK_HEAD; 164; 1.

FT PROSITE PS50039; FORK_HEAD; 165; 1.

FT PROSITE PS50039; FORK_HEAD; 166; 1.

FT PROSITE PS50039; FORK_HEAD; 167; 1.

FT PROSITE PS50039; FORK_HEAD; 168; 1.

FT PROSITE PS50039; FORK_HEAD; 169; 1.

FT PROSITE PS50039; FORK_HEAD; 170; 1.

FT PROSITE PS50039; FORK_HEAD; 171; 1.

FT PROSITE PS50039; FORK_HEAD; 172; 1.

FT PROSITE PS50039; FORK_HEAD; 173; 1.

FT PROSITE PS50039; FORK_HEAD; 174; 1.

FT PROSITE PS50039; FORK_HEAD; 175; 1.

FT PROSITE PS50039; FORK_HEAD; 176; 1.

FT PROSITE PS50039; FORK_HEAD; 177; 1.

FT PROSITE PS50039; FORK_HEAD; 178; 1.

FT PROSITE PS50039; FORK_HEAD; 179; 1.

FT PROSITE PS50039; FORK_HEAD; 180; 1.

FT PROSITE PS50039; FORK_HEAD; 181; 1.

FT PROSITE PS50039; FORK_HEAD; 182; 1.

FT PROSITE PS50039; FORK_HEAD; 183; 1.

FT PROSITE PS50039; FORK_HEAD; 184; 1.

FT PROSITE PS50039; FORK_HEAD; 185; 1.

FT PROSITE PS50039; FORK_HEAD; 186; 1.

FT PROSITE PS50039; FORK_HEAD; 187; 1.

FT PROSITE PS50039; FORK_HEAD; 188; 1.

FT PROSITE PS50039; FORK_HEAD; 189; 1.

FT PROSITE PS50039; FORK_HEAD; 190; 1.

FT PROSITE PS50039; FORK_HEAD; 191; 1.

FT PROSITE PS50039; FORK_HEAD; 192; 1.

FT PROSITE PS50039; FORK_HEAD; 193; 1.

FT PROSITE PS50039; FORK_HEAD; 194; 1.

FT PROSITE PS50039; FORK_HEAD; 195; 1.

FT PROSITE PS50039; FORK_HEAD; 196; 1.

FT PROSITE PS50039; FORK_HEAD; 197; 1.

FT PROSITE PS50039; FORK_HEAD; 198; 1.

FT PROSITE PS50039; FORK_HEAD; 199; 1.

FT PROSITE PS50039; FORK_HEAD; 200; 1.

FT PROSITE PS50039; FORK_HEAD; 201; 1.

FT PROSITE PS50039; FORK_HEAD; 202; 1.

FT PROSITE PS50039; FORK_HEAD; 203; 1.

FT PROSITE PS50039; FORK_HEAD; 204; 1.

FT PROSITE PS50039; FORK_HEAD; 205; 1.

FT PROSITE PS50039; FORK_HEAD; 206; 1.

FT PROSITE PS50039; FORK_HEAD; 207; 1.

FT PROSITE PS50039; FORK_HEAD; 208; 1.

FT PROSITE PS50039; FORK_HEAD; 209; 1.

FT PROSITE PS50039; FORK_HEAD; 210; 1.

FT PROSITE PS50039; FORK_HEAD; 211; 1.

FT PROSITE PS50039; FORK_HEAD; 212; 1.

FT PROSITE PS50039; FORK_HEAD; 213; 1.

FT PROSITE PS50039; FORK_HEAD; 214; 1.

FT PROSITE PS50039; FORK_HEAD; 215; 1.

FT PROSITE PS50039; FORK_HEAD; 216; 1.

FT PROSITE PS50039; FORK_HEAD; 217; 1.

FT PROSITE PS50039; FORK_HEAD; 218; 1.

FT PROSITE PS50039; FORK_HEAD; 219; 1.

FT PROSITE PS50039; FORK_HEAD; 220; 1.

FT PROSITE PS50039; FORK_HEAD; 221; 1.

FT PROSITE PS50039; FORK_HEAD; 222; 1.

FT PROSITE PS50039; FORK_HEAD; 223; 1.

FT PROSITE PS50039; FORK_HEAD; 224; 1.

FT PROSITE PS50039; FORK_HEAD; 225; 1.

FT PROSITE PS50039; FORK_HEAD; 226; 1.

FT PROSITE PS50039; FORK_HEAD; 227; 1.

FT PROSITE PS50039; FORK_HEAD; 228; 1.

FT PROSITE PS50039; FORK_HEAD; 229; 1.

FT PROSITE PS50039; FORK_HEAD; 230; 1.

FT PROSITE PS50039; FORK_HEAD; 231; 1.

FT PROSITE PS50039; FORK_HEAD; 232; 1.

FT PROSITE PS50039; FORK_HEAD; 233; 1.

FT PROSITE PS50039; FORK_HEAD; 234; 1.

FT PROSITE PS50039; FORK_HEAD; 235; 1.

FT PROSITE PS50039; FORK_HEAD; 236; 1.

FT PROSITE PS50039; FORK_HEAD; 237; 1.

FT PROSITE PS50039; FORK_HEAD; 238; 1.

FT PROSITE PS50039; FORK_HEAD; 239; 1.

FT PROSITE PS50039; FORK_HEAD; 240; 1.

FT PROSITE PS50039; FORK_HEAD; 241; 1.

FT PROSITE PS50039; FORK_HEAD; 242; 1.

FT PROSITE PS50039; FORK_HEAD; 243; 1.

FT PROSITE PS50039; FORK_HEAD; 244; 1.

FT PROSITE PS50039; FORK_HEAD; 245; 1.

FT PROSITE PS50039; FORK_HEAD; 246; 1.

FT PROSITE PS50039; FORK_HEAD; 247; 1.

FT PROSITE PS50039; FORK_HEAD; 248; 1.

FT PROSITE PS50039; FORK_HEAD; 249; 1.

FT PROSITE PS50039; FORK_HEAD; 250; 1.

FT PROSITE PS50039; FORK_HEAD; 251; 1.

FT PROSITE PS50039; FORK_HEAD; 252; 1.

FT PROSITE PS50039; FORK_HEAD; 253; 1.

FT PROSITE PS50039; FORK_HEAD; 254; 1.

FT PROSITE PS50039; FORK_HEAD; 255; 1.

FT PROSITE PS50039; FORK_HEAD; 256; 1.

FT PROSITE PS50039; FORK_HEAD; 257; 1.

FT PROSITE PS50039; FORK_HEAD; 258; 1.

FT PROSITE PS50039; FORK_HEAD; 259; 1.

FT PROSITE PS50039; FORK_HEAD; 260; 1.

FT PROSITE PS50039; FORK_HEAD; 261; 1.

FT PROSITE PS50039; FORK_HEAD; 262; 1.

FT PROSITE PS50039; FORK_HEAD; 263; 1.

FT PROSITE PS50039; FORK_HEAD; 264; 1.

FT PROSITE PS50039; FORK_HEAD; 265; 1.

FT PROSITE PS50039; FORK_HEAD; 266; 1.

FT PROSITE PS50039; FORK_HEAD; 267; 1.

FT PROSITE PS50039; FORK_HEAD; 268; 1.

FT PROSITE PS50039; FORK_HEAD; 269; 1.

FT PROSITE PS50039; FORK_HEAD; 270; 1.

FT PROSITE PS50039; FORK_HEAD; 271; 1.

FT PROSITE PS50039; FORK_HEAD; 272; 1.

FT PROSITE PS50039; FORK_HEAD; 273; 1.

FT PROSITE PS50039; FORK_HEAD; 274; 1.

FT PROSITE PS50039; FORK_HEAD; 275; 1.

FT PROSITE PS50039; FORK_HEAD; 276; 1.

FT PROSITE PS50039; FORK_HEAD; 277; 1.

FT PROSITE PS50039; FORK_HEAD; 278; 1.

FT PROSITE PS50039; FORK_HEAD; 279; 1.

FT PROSITE PS50039; FORK_HEAD; 280; 1.

FT PROSITE PS50039; FORK_HEAD; 281; 1.

FT PROSITE PS50039; FORK_HEAD; 282; 1.

FT PROSITE PS50039; FORK_HEAD; 283; 1.

FT PROSITE PS50039; FORK_HEAD; 284; 1.

FT PROSITE PS50039; FORK_HEAD; 285; 1.

FT PROSITE PS50039; FORK_HEAD; 286; 1.

FT PROSITE PS50039; FORK_HEAD; 287; 1.

FT PROSITE PS50039; FORK_HEAD; 288; 1.

FT PROSITE PS50039; FORK_HEAD; 289; 1.

FT PROSITE PS50039; FORK_HEAD; 290; 1.

FT PROSITE PS50039; FORK_HEAD; 291; 1.

FT PROSITE PS50039; FORK_HEAD; 292; 1.

FT PROSITE PS50039; FORK_HEAD; 293; 1.

FT PROSITE PS50039; FORK_HEAD; 294; 1.

FT PROSITE PS50039; FORK_HEAD; 295; 1.

FT PROSITE PS50039; FORK_HEAD; 296; 1.

FT PROSITE PS50039; FORK_HEAD; 297; 1.

FT PROSITE PS50039; FORK_HEAD; 298; 1.

FT PROSITE PS50039; FORK_HEAD; 299; 1.

FT PROSITE PS50039; FORK_HEAD; 300; 1.

FT PROSITE PS50039; FORK_HEAD; 301; 1.

FT PROSITE PS50039; FORK_HEAD; 302; 1.

FT PROSITE PS50039; FORK_HEAD; 303; 1.

FT PROSITE PS50039; FORK_HEAD; 304; 1.

FT PROSITE PS50039; FORK_HEAD; 305; 1.

FT PROSITE PS50039; FORK_HEAD; 306; 1.

FT PROSITE PS50039; FORK_HEAD; 307; 1.

FT PROSITE PS50039; FORK_HEAD; 308; 1.

FT PROSITE PS50039; FORK_HEAD; 309; 1.

FT PROSITE PS50039; FORK_HEAD; 310; 1.

FT PROSITE PS50039; FORK_HEAD; 311; 1.

FT PROSITE PS50039; FORK_HEAD; 312; 1.

FT PROSITE PS50039; FORK_HEAD; 313; 1.

FT PROSITE PS50039; FORK_HEAD; 314; 1.

FT PROSITE PS50039; FORK_HEAD; 315; 1.

FT PROSITE PS50039; FORK_HEAD; 316; 1.

FT PROSITE PS50039; FORK_HEAD; 317; 1.

FT PROSITE PS50039; FORK_HEAD; 318; 1.

FT PROSITE PS50039; FORK_HEAD; 319; 1.

FT PROSITE PS50039; FORK_HEAD; 320; 1.

FT PROSITE PS50039; FORK_HEAD; 321; 1.

FT PROSITE PS50039; FORK_HEAD; 322; 1.

FT PROSITE PS50039; FORK_HEAD; 323; 1.

FT PROSITE PS50039; FORK_HEAD; 324; 1.

FT PROSITE PS50039; FORK_HEAD; 325; 1.

FT PROSITE PS50039; FORK_HEAD; 326; 1.

FT PROSITE PS50039; FORK_HEAD; 327; 1.

FT PROSITE PS50039; FORK_HEAD; 328; 1.

FT PROSITE PS50039; FORK_HEAD; 329; 1.

FT PROSITE PS50039; FORK_HEAD; 330; 1.

FT PROSITE PS50039; FORK_HEAD; 331; 1.

FT PROSITE PS50039; FORK_HEAD; 332; 1.

FT PROSITE PS50039; FORK_HEAD; 333; 1.

FT PROSITE PS50039; FORK_HEAD; 334; 1.

FT PROSITE PS50039; FORK_HEAD; 335; 1.

FT PROSITE PS50039; FORK_HEAD; 336; 1.

FT PROSITE PS50039; FORK_HEAD; 337; 1.

FT PROSITE PS50039; FORK_HEAD; 338; 1.

FT PROSITE PS50039; FORK_HEAD; 339; 1.

FT PROSITE PS50039; FORK_HEAD; 340; 1.

FT PROSITE PS50039; FORK_HEAD; 341; 1.

FT PROSITE PS50039; FORK_HEAD; 342; 1.

FT PROSITE PS50039; FORK_HEAD; 343; 1.

FT PROSITE PS50039; FORK_HEAD; 344; 1.

FT PROSITE PS50039; FORK_HEAD; 345; 1.

FT PROSITE PS50039; FORK_HEAD; 346; 1.

FT PROSITE PS50039; FORK_HEAD; 347; 1.

FT PROSITE PS50039; FORK_HEAD; 348; 1.

FT PROSITE PS50039; FORK_HEAD; 349; 1.

FT PROSITE PS50039; FORK_HEAD; 350; 1.

FT PROSITE PS50039; FORK_HEAD; 351; 1.

FT PROSITE PS50039; FORK_HEAD; 352; 1.

FT PROSITE PS50039; FORK_HEAD; 353; 1.

FT PROSITE PS50039; FORK_HEAD; 354; 1.

FT PROSITE PS50039; FORK_HEAD; 355; 1.

FT PROSITE PS50039; FORK_HEAD; 356; 1.

FT PROSITE PS50039; FORK_HEAD; 357; 1.

FT PROSITE PS50039; FORK_HEAD; 358; 1.

FT PROSITE PS50039; FORK_HEAD; 359; 1.

FT PROSITE PS50039; FORK_HEAD; 360; 1.

FT PROSITE PS50039; FORK_HEAD; 361; 1.

FT PROSITE PS50039; FORK_HEAD; 362; 1.

FT PROSITE PS50039; FORK_HEAD; 363; 1.

FT PROSITE PS50039; FORK_HEAD; 364; 1.

FT PROSITE PS50039; FORK_HEAD; 365; 1.

FT PROSITE PS50039; FORK_HEAD; 366; 1.

FT PROSITE PS50039; FORK_HEAD; 367; 1.

FT PROSITE PS50039; FORK_HEAD; 368; 1.

FT PROSITE PS50039; FORK_HEAD; 369; 1.

FT PROSITE PS50039; FORK_HEAD; 370; 1.

FT PROSITE PS50039; FORK_HEAD; 371; 1.

FT PROSITE PS50039; FORK_HEAD; 372; 1.

FT PROSITE PS50039; FORK_HEAD; 373; 1.

FT PROSITE PS50039; FORK_HEAD; 374; 1.

FT PROSITE PS50039; FORK_HEAD; 375; 1.

FT PROSITE PS50039; FORK_HEAD; 376; 1.

FT PROSITE PS50039; FORK_HEAD; 377; 1.

FT PROSITE PS50039; FORK_HEAD; 378; 1.

FT PROSITE PS50039; FORK_HEAD; 379; 1.

FT PROSITE PS50039; FORK_HEAD; 380; 1.

FT PROSITE PS50039; FOR

the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to licensee@isb-sib.ch).

PRODOM; PD000425; TF_Fork_head; 1.
SMART; SN00339; FH_1.
PROSITE; PS00657; FORK_HEAD_1; FALSE_NEG.
PROSITE; PS00558; FORK_HEAD_2; 1.
PROSITE; PS50039; FORK_HEAD_3; 1.
Transcription regulation; DNA-binding; Nuclear protein; Activator;
DNA_BIND 66 143
DOMAIN 266 269
DOMAIN 290 296
DOMAIN 306 314
SEQUENCE 565 AA; 61569 MW; 917BAFF3F9227AD4 CRC64;

Query Match Score 201.5; DB 1; Length 565;
Best Local Similarity 44.4%; Pred. No. 1..1e-14;
Matches 40; Conservative 18; Mismatches 23; Indels 9; Gaps 3;

14 SYAELITTAIMASPEKRLTLAGYYENWNQNVPYFRDKGDSNSAGWKNSIRHNLSHSRF 73
70 SYATLTYAINDSPAKXMTLSIYWRNCIDNPYYK----NAGIGWKNSIRHNLSLNKCF 124

74 MRI--QNEGAKGSWWVINEPAKPGMNPQR 101
125 RKVPRPDDPGKGSWYI--DTCPDISRK 152

RESULT 8
ID FXJ2_HUMAN
Q9PK8; Q96PS9; Q9NSNS;
STANDARD;
PRT; 574 AA.
16-OCT-2001 (Rel. 40, Created)
16-OCT-2001 (Rel. 40, Last sequence update)
10-OCT-2003 (Rel. 42, Last annotation update)
Foxhead box protein J2 (Fork head homologous X).
FOXJ2 OR FHx.
Homo sapiens (Human).
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
NCBI_TAXID=9606;
[1] SEQUENCE FROM N.A. (ISOFORM FOXJ2.L).
MEDLINE=20239944; Published=10/77590;
Perez-Sanchez C.; Gomez-Ferreria M.A.; de la Fuente C.A.,
Granadino B.; Velasco G.; Esteban A.; Rey-Campos J.;
"FHx, a novel fork head factor with a dual DNA binding specificity.";
J. Biol. Chem. 275:12909-12916 (2000).
[2] SEQUENCE FROM N.A. (ISOFORM FOXJ2.S).
MEDLINE=20425082; Published=10/66786;
Perez-Sanchez C.; de la Fuente C.A., Gomez-Ferreria M.A.,
Granadino B.; Rey-Campos J.;
"FHx, two isoforms of the human fork-head factor FHx
(FoxJ2) with different activity.";
J. Mol. Biol. 301:1795-1806 (2000).
[3] SEQUENCE OF 458-574 FROM N.A. (ISOFORM FOXJ2.L).
Tissue=melanoma;
Ansorge W.; Wirkner U.; Mewes H.-W.; Weil B.; Wiemann S.;
Submitted (MAR-2000) to the EMBL/GenBank/DDBJ databases.
!- FUNCTION: transcriptional activator. Able to bind to two different
type of DNA binding sites. Isoform FOXJ2.L behaves as a more
potent transactivator than FOXJ2.S.
!- SUBCELLULAR LOCATION: Nuclear.
!- ALTERNATIVE PRODUCTS:
Event=Alternative splicing; Named isoforms=2;
Name=FOXJ2.L; Synonyms=FOX.L;
IsoID=Q9PK8; Sequence=Display;
Name=FOXJ2_S; Synonyms=FOX.S;
IsoID=Q9PK8_2; Sequence=VSP_001544;
!- TISSUE SPECIFICITY: Widely expressed.
!- SIMILARITY: Contains 1 fork-head domain.

RESULT 9
FXJ1_HUMAN
ID FXJ1_HUMAN
STANDARD;
PRT; 351 AA.
AC Q14518; Q9N6L8;
DT 01-NOV-1997 (Rel. 35, Created)
DT 30-MAR-2000 (Rel. 39, Last sequence update)
DT 15-MAR-2004 (Rel. 43, Last annotation update)
DB Forkhead box Protein II1 (Forkhead-related protein FKHL10) (Forkhead-related transcription factor 6) (FBXAC-6) (Hepatocyte nuclear factor 3
DE forktail homolog 3) (HNF3/fork-head homolog-3) (HFFH-3).
GN Homo sapiens (Human).
OS Homo sapiens.
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TAXID=9606;
RN [1]
RP SEQUENCE FROM N.A. (ISOFORM 1).
RC TISSUE=Kidney;
RC MEDLINE=97298077; Published=951225;
RX Overdier D.G.; Ye H.; Peterson R.S.; Clevidence D.E.; Costa R.H.;
RA "The winged helix transcriptional activator HFFH-3 is expressed in the
distal tubules of embryonic and adult mouse kidney."
RT CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC [2] J. Biol. Chem. 272:13725-13730 (1997).
RN [2]

SEQUENCE FROM N.A. (ISOFORM 2).	DR	PROSITE; PS50039; FORK_HEAD_3; 1
RC TISSUE-Colon, and Kidney;	KW	DNA-binding; Nuclear protein; Transcription regulation; Activator;
BX MEDLINE=23388257; Published=12/27/932;	KW	Alternative splicing.
RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,	KW	FORK-HEAD.
RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,	FT	Missing (in isoform 2).
RA Altschul S.F., Zeeberg B., Buetow K.H., Schaeffer B.C.F., Bhat N.K.,	FT	VARSPPLIC 186
RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,	FT	VARSPPLIC 259
RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,	FT	/FTID=VSP 001543.
RA Brownstein M.J., Soresi M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,	SQ	SEQUENCE 351 AA; 38091 MW; 31198DC2003913D51 CRC64;
RA Baha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullally S.J.,	Query Match	Score 190; DB 1; Length 351;
RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,	Best Local Similarity	40.0%; Pred. No. 1.2e-13;
RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,	Matches	38; Conservative 22; Mismatches 25; Indels 10; Gaps 3
RA Villalon D.K., Muzny D.M., Sodergren B.J., Lu X., Bouffard G.G.,	Qy	14 SYAELITTAIMASPERKUTLAQYEWWMQNVYPFRDGDSNSGAWNSIRNLNSLHSRF 73
RA Fahey J., Helton E., Kettenm M., Macan A., Rodrigues S., Sanchez A.,	Db	100 SYASALIAMAHGAPDKRLTSQLQYQVADNFPFY ---NRSKGWNSTRHNLNDCF 154
RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,	Qy	74 MRI-QNEGAGATGATINPDAX---PGMNPRTTR 103
RA Blakesley R.W., Touchman J.W., Green B.D., Dickson M.C.,	Db	155 KKVRPDRDDPGKGNWTLDPNCEKMFQDNGNFRKR 189
RA Rodriguez A.C., Grimwade J., Butterfield Y.N., Krzywinski M.I., Skalska U., Smailus D.E.,		
RA Schnarch A., Schein J.E., Jones S.J.M., Marra M.A.,		
RT Generation and initial analysis of more than 15,000 full-length		
RT human and mouse cDNA sequences".		
RL Proc. Natl. Acad. Sci. U.S.A. 99:16399-16403 (2002).		
[3]		
SEQUENCE OF 91-166 FROM N.A.	RESULT 1.0	
RX MEDLINE=95045392; PubMed=7957066;	ID YK98_SC9PO STANDARD; PRT; 743 AA.	
PA Pierrou S., Heiligquist M., Samuelsson L., Enerbaeck S., Carlsson P.;	AC O14270; Q9P7F6;	
RT "Cloning and characterization of seven human forkhead proteins:	DT 28-FEB-2003 (Rel. 41, Created)	
RT binding site specificity and DNA bending.";	DT 28-FEB-2003 (Rel. 41, Last sequence update)	
EMBO J. 13:5002-5012(1994).	DT 28-FEB-2003 (Rel. 41, Last annotation update)	
RL J.-I. FUNCTION: Transcription activator. May act on the genes for Na/K-ATPase, Na/H and anion exchangers, E-cadherin and mineralocorticoid receptor as well as on the genes for transcription factors HNF-1, VHNF-1 and HNF-4.	DT Putative forkhead-related transcription factor C1142.08.	
CC SUBCELLULAR LOCATION: Nuclear.	GN SPAC1142.08 OR SPACC9.01.	
CC ATPase, Na/H and anion exchangers, E-cadherin and mineralocorticoid receptor as well as on the genes for transcription factors HNF-1, VHNF-1 and HNF-4.	OS Schizosaccharomyces pombe (Fission Yeast).	
CC SUBCELLULAR LOCATION: Nuclear.	OC Schizosaccharomyces pombe; Ascomycota; Schizosaccharomyces pombe; Schizosaccharomyces pombe; Schizosaccharomyces pombe; Schizosaccharomyces pombe.	
CC ALTERNATIVE PRODUCTS:	OC Schizosaccharomyces pombe.	
Name=?;	OX NCBI_TAXID=4896;	
CC Event-Alternative splicing; Named isoforms=2;	RN [1]	
CC IsoId=Q12951-1; Sequence=Displayed;	RP SEQUENCE FROM N.A.	
CC Name=?;	RC STRAIN=9/2;	
CC IsoId=Q12951-2; Sequence=VSP 001543;	RA MEDLINE=21848401; PubMed=11859360;	
CC Note=No experimental confirmation available;	RA Wood V., Gwilliam R., Rajandream M.A., Lyne M., Lyne R., Stewart A.,	
CC !- TISSUE-SPECIFICITY: Kidney specific;	RA Sgourou S., Peat N., Hayles J., Baker S., Basham D., Bowman S.,	
CC !- SIMILARITY: Contains 1 fork-head domain.	RA Brooks K., Brown D., Brown S., Chillingworth T., Churcher C.M.,	
This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the FMBL outstation - use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See http://www.isb-sib.ch/announce/ or send an email to license@isb-sib.ch).	RA Collins M., Connor R., Cronin A., Davis P., Feltwell T., Fraser A.,	
EMBL; Li13203; AAB0574.1; -	RA Gentles S., Goble A., Hamlin N., Harris D., Hidalgo J., Hodgson G.,	
EMBL; BC029778; AAH29778.1; -	RA James K., Jones L., Jones M., Leather S., McDonald S., McLean J.,	
EMBL; U13224; AAA2041.1; -	RA Mooney P., Moul S., Margall C., Murphy L., Niblett D., Odell C.,	
EMBL; S51629; S51629.1; -	RA Oliver K., O'Neil S., Pearson D., Quail M.A., Rabbinowitz E.,	
HSSP; Q63245; 2HFF.	RA Rutherford K., Rutter S., Saunders D., Seeger K., Sharp S.,	
TRANSFAC; T0447; -	RA Skelton J., Simmonds M., Squares R., Stevens K., Taylor K., Taylor R.G., Tivey A., Walsh S.V., Warren T., Whitehead S.,	
Genew; HGNC:3815; FOX11.	RA Woodward J., Volckaert G., Aert R., Robben J., Grymonpre B.,	
MIM; 601093; -	RA Weltjens I., Vansstreets E., Rieger M., Schaefer M., Mueller-Auer S.,	
GO; GO:0005634; C:nucleus; NAS.	RA Fuchs M., Fritz C., Holzer B., Moestl D., Hilbert H.,	
GO; GO:000830; F:DNA binding activity; NAS.	RA Domínguez A., Revuelta J.L., Sanchez M., del Rey F., Benito J.,	
GO; GO:000370; F:transcription factor activity; TAS.	RA Cerruti L., Lowe T., McCombie W.R., Paulsen I., Potashkin J.,	
GO; GO:0007275; P:transcription, DNA-dependent; NAS.	RA Supakowski G.V., Usseyry D., Barrell B.G., Nurse P.,	
InterPro; IPR01766; TF_Fork_head.	RT RT "The genome sequence of Schizosaccharomyces pombe.";	
PRINTS; PP00025; Fork_head; 1.	RL Nature 415:871-880 (2002).	
PRODOM; PD000425; TF_Fork_head; 1.	CC CC !- SIMILARITY: Contains 1 fork-head domain.	
SMART; SM00339; FH; 1.	CC CC !- SIMILARITY: Contains 1 FHA domain.	
PROSITE; PS00657; FORK_HEAD_1; 1.		
PROSITE; PS00658; FORK_HEAD_2; 1.		

entities requires a license agreement (see <http://www.isb-sib.ch/announce/> or send an email to license@isb-sib.ch).

CC	DR EMBL; L10719; ; NOT_ANNOTATED_CDS.
CC	DR HSSP; Q63245; 2HFF.
CC	DR TRANSPAC; T01831; -.
CC	DR InterPro; IPR01766; TP_Fork_head.
CC	DR DRINTERPRO; P00050; Fork_head; 1.
CC	DR PFAM; PF00050; Fork_head; 1.
CC	DR PRINTS; PRO0053; FORKHEAD.
CC	DR PRODOM; PD000425; TF_Fork_head; 1.
CC	DR SMART; SM00339; FH; 1.
CC	DR PROSITE; PS00657; FORK_HEAD 1; 1.
CC	DR PROSITE; PS00658; FORK_HEAD 2; 1.
CC	DR PROSITE; PS00039; FORK_HEAD 3; 1.
CC	DR DNA-binding; Nuclear protein; Oncogene.
CC	FT DOMAIN 42 54 POLY-HTS.
CC	FT DOMAIN 55 58 POLY-ALA.
CC	FT DOMAIN 64 67 POLY-ALA.
CC	FT DOMAIN 101 106 POLY-ALA.
CC	FT DNA BIND 142 233 FORK-HEAD.
CC	SQ SEQUENCE 387 AA: 42283 MW; FEAF902F50FFB42F9 CRC64;
CC	Query Match 34.0%; Score 188.5%; DB 1; Length 387;
CC	Best Local Similarity 43.9%; Pred. No. 1.9e-13; Indels 7; Gaps 2;
CC	Matches 36; Conservative 20; Mismatches 19;
CC	QY 14 SYAPLITTAIMASPEKRLLAQYEWMMVNPYFRDKGDSNSAGWKNSIREHNLSHSRF 73
CC	Db 147 SYNLLIMMAIRQSQEKRLLTNGIYEFINKNPFYRE----NFCQGMQNSIREHNLSNRCF 201
CC	QY 74 MRQIN-EGAGKSSWWVNPDAA 93
CC	Db 202 VKVPHYDDGGKGNiWMLDSS 223
CC	RESULT 12
CC	FXGB_CHICK STANDARD; PRT; 451 AA.
CC	ID FXGB_CHICK
CC	AC Q9D9E4;
CC	DT 15-JUL-1998 (Rel. 36, Created)
CC	DT 15-JUL-1998 (Rel. 36, Last sequence update)
CC	DT 15-MAR-2004 (Rel. 43, Last annotation update)
CC	DE Forkhead box protein G1B (Forkhead-related protein FKH1)
CC	DE (Transcription factor BF-1) (Brain factor 11) (BF1) (Proto-
CC	DE oncogene C-QIN) (N-62-5) (CBQ 3-1).
CC	GN FOXG1B OR FKH1 OR QIN.
CC	OS Gallus gallus (Chicken).
CC	OC Bivalvia; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
CC	OC Archosauria; Aves; Neognathae; Galliformes; Phasianidae; Phasianinae;
CC	OC Gallus.
CC	OC NCBITaxonID=9031;
CC	OX
CC	RN [1]
CC	SEQUENCE FROM N.A.
CC	RX MEDLINE=95132616; PubMed=7831308;
CC	RA Chang H.W.; Li J.; Kretzschmar D.; Vogt P.K.;
CC	RT "Avian cellular homolog of the qin oncogene.";
CC	RL Proc. Natl. Acad. Sci. U.S.A. 92:447-451(1995).
CC	RN [2]
CC	RP SEQUENCE FROM N.A.
CC	RC STRAIN=White Leghorn; TISSUE=Retina;
CC	RC MEDLINE=96338226; PubMed=8757134;
CC	RA Yuasa J.; Hirano S.; Yamagata M.; Noda M.;
CC	RT "Visual projection maps specified by topographic expression of transcription factors in the retina."
CC	RL Nature 382:632-635 (1996).
CC	-!- FUNCTION: May determine the nasotemporal axis of the retina, and consequently the topographical projection of the retinal ganglion-cell axons to the tectum by controlling expression of their target genes.
CC	CC -!- SUBCELLULAR LOCATION: Nuclear (Potential).
CC	CC -!- TISSUE_SPECIFICITY: Retina and brain.
CC	CC -!- DEVELOPMENTAL_STAGE: Can be detected in regions including primordial retina and neuroepithelium by embryonic day 2 (E2). At E3, expressed in the nasal retina and pigment epithelium as well as in the telencephalon, and at E7 is expressed in retinal ganglion cells. Levels begin to decline from E4 and almost
CC	CC -!- SIMILARITY: Contains 1 fork-head domain.
CC	CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use. By non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See http://www.isb-sib.ch/announce/ or send an email to license@isb-sib.ch).
CC	CC

CC disappear by E10.
 CC !- SIMILARITY: Contains 1 fork-head domain.
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC
 DR EMBL; L3814; AAA6954.1; -;
 DR EMBL; U7275; AAB8466.1; -;
 DR PTB; A55909; A55909.
 DR HSP; Q63245; 2HFF.
 DR TRSFAC; T01833; -;
 DR InterPro; IPR001766; TF_Fork_head.
 DR Pfam; PF00250; Fork_head; 1.
 DR PRINTS; PR00053; FORKHEAD.
 DR ProdDom; PD000425; TF_Fork_head; 1.
 DR SMART; SM00339; FH; 1.
 DR PROSITE; PS00657; FORK_HEAD_1; 1.
 DR PROSITE; PS00658; FORK_HEAD_2; 1.
 DR PROSITE; PS50039; FORK_HEAD_3; 1.
 KW Transcription regulation; DNA-binding; Nuclear protein;
 DOMAIN 45 POLY-HIS.
 FT DOMAIN 55 58
 FT DOMAIN 64 67
 FT DOMAIN 103 106
 FT DNA_BIND 142 233
 SQ SEQUENCE 451 AA; 48856 MW; E9E5B407D2321B50 CRC64;
 Qy 14 SYAELTTAIMASPEKRLTIAQVYEMVNQVNPYFRDKGDSNSAGWNKNSIRHNLSLHSRF 73
 Best Local Similarity 43.9%; Pred. No. 2.3e-13; Mismatches 19; Indels 7; Gaps 2;
 Matches 36; Conservative 20; Mismatches 19; Indels 7; Gaps 2;
 Db 147 SYNALIMMAIROSPEKRLTIAQVYEMVNQVNPYFRDKGDSNSAGWNKNSIRHNLSLHSRF 201
 Qy 74 MRIQN-EGAGKSSWWVINDPA 93
 Db 202 VKVPRHYDDPGKGNVWMLDPSS 223
 Qy 14 SYAELTTAIMASPEKRLTIAQVYEMVNQVNPYFRDKGDSNSAGWNKNSIRHNLSLHSRF 73
 Best Local Similarity 43.9%; Pred. No. 2.3e-13; Mismatches 19; Indels 7; Gaps 2;
 Matches 36; Conservative 20; Mismatches 19; Indels 7; Gaps 2;
 Db 147 SYNALIMMAIROSPEKRLTIAQVYEMVNQVNPYFRDKGDSNSAGWNKNSIRHNLSLHSRF 201
 Qy 74 MRIQN-EGAGKSSWWVINDPA 93
 Db 228 VKVPRHYDDPGKGNVWMLDPSS 249

RESULT 14

ID	FXGB RAT	STANDARD	PRT	480 AA.
AC	Q00539;			
DT	01-JUL-1993	(Rel. 26, Created)		
DT	01-OCT-2001	(Rel. 40, Last sequence update)		
DE	Forhead box protein G1B (Porkhead-related protein FKH1)			
DE	(Transcription factor BF-1) (Brain factor 1) (BF1).			
GN	FOXGB OR FKH1.			
OS	Homo sapiens (Human).			
OC	Eukaryota; Chordata; Craniata; Vertebrata; Euteleostomi;			
OC	Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.			
OX	[1] _TaxID=9606;			
RN				
RP	SEQUENCE FROM N.A.			
RC	SEQUENCE=Brain;			
RX	Medline=3504832; PubMed=7959731;			
RA	Murphy D.B., Wiese S., Burfeind P., Schmundt D., Mattei M.-G.,			
RA	Schulz-Schaeffer W., Thies U.,			
RT	"Human brain factor 1, a new member of the fork head gene family.";			
RL	Genomics 21:551-557 (1994).			
RN	[2]			
RP	SEQUENCE FROM N.A.			
RC	SEQUENCE=Brain;			
RX	Medline=5322450; PubMed=7599184;			
RA	Wiese S., Murphy D.B., Burfeind P., Schmundt D.,			

RESULT 13

ID	FXGB RAT	STANDARD	PRT	480 AA.
AC	Q00539;			
DT	01-JUL-1993	(Rel. 26, Created)		
DT	16-OCT-2001	(Rel. 40, Last sequence update)		
DE	Forhead box protein G1B (Porkhead-related protein FKH1)			
DE	(Transcription factor BF-1) (Brain factor 1) (BF1).			
GN	FOXGB OR FKH1.			
OS	Rattus norvegicus (Rat).			
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;			
OC	Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.			
OX	[1] _TaxID=10116;			
RN				
RP	SEQUENCE FROM N.A.			
RC	SEQUENCE=Brain.			

RA RT "The genes for human brain factor 1 and 2, members of the fork head gene family, are clustered on chromosome 14q.";
 RA RBL Biochim Biophys Acta 1262:105-112 (1995).
 CC !- FUNCTION: Plays an important role in the establishment of the regional subdivision of the developing brain and in the development of the telencephalon. Sequence-specific DNA-binding protein with a distinct binding specificity (By similarity).
 CC !- SUBCELLULAR LOCATION: Nuclear.
 CC !- SIMILARITY: Contains 1 fork-head domain.
 CC
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to license@isb-sib.ch).
 CC
 DR EMBL; X74142; CAA52239.1; -;
 DR HSSP; Q63245; 2HFF.
 DR TRANSFAC; T023350; -;
 DR Genev; HGNC;3112; FOXGB.
 DR GO; GO-0003677; F:DNA binding; TAS.
 DR GO; GO-0007420; P:brain development; TAS.
 DR MIM; 164874; -;
 DR InterPro; IPR01766; TF_Fork_head.
 DR Pfam; PF00250; Fork_head; 1.
 DR PRINTS; PR00053; FORKHEAD.
 DR ProdDom; PD000425; TF_Fork_head; 1.
 DR SMART; SM00339; FH; 1.
 DR PROSITE; PS00657; FORK_HEAD_1; 1.
 DR PROSITE; PS00658; FORK_HEAD_2; 1.
 DR PROSITE; PS50039; FORK_HEAD_3; 1.
 DR PS00659; FORK_HEAD_4; 1.
 DR PROSITE; PS50040; FORK_HEAD_5; 1.
 DR PROSITE; PS50041; FORK_HEAD_6; 1.
 DR PROSITE; PS50042; FORK_HEAD_7; 1.
 DR PROSITE; PS50043; FORK_HEAD_8; 1.
 DR PROSITE; PS50044; FORK_HEAD_9; 1.
 DR PROSITE; PS50045; FORK_HEAD_10; 1.
 DR PROSITE; PS50046; FORK_HEAD_11; 1.
 DR PROSITE; PS50047; FORK_HEAD_12; 1.
 DR PROSITE; PS50048; FORK_HEAD_13; 1.
 DR PROSITE; PS50049; FORK_HEAD_14; 1.
 DR PROSITE; PS50050; FORK_HEAD_15; 1.
 DR PROSITE; PS50051; FORK_HEAD_16; 1.
 DR PROSITE; PS50052; FORK_HEAD_17; 1.
 DR PROSITE; PS50053; FORKHEAD.
 DR PROSITE; PS50054; FORK_HEAD_18; 1.
 DR PROSITE; PS50055; FORK_HEAD_19; 1.
 DR PROSITE; PS50056; FORK_HEAD_20; 1.
 DR PROSITE; PS50057; FORK_HEAD_21; 1.
 DR PROSITE; PS50058; FORK_HEAD_22; 1.
 DR PROSITE; PS50059; FORK_HEAD_23; 1.
 DR PROSITE; PS50060; FORK_HEAD_24; 1.
 DR PROSITE; PS50061; FORK_HEAD_25; 1.
 DR PROSITE; PS50062; FORK_HEAD_26; 1.
 DR PROSITE; PS50063; FORK_HEAD_27; 1.
 DR PROSITE; PS50064; FORK_HEAD_28; 1.
 DR PROSITE; PS50065; FORK_HEAD_29; 1.
 DR PROSITE; PS50066; FORK_HEAD_30; 1.
 DR PROSITE; PS50067; FORK_HEAD_31; 1.
 DR PROSITE; PS50068; FORK_HEAD_32; 1.
 DR PROSITE; PS50069; FORK_HEAD_33; 1.
 DR PROSITE; PS50070; FORK_HEAD_34; 1.
 DR PROSITE; PS50071; FORK_HEAD_35; 1.
 DR PROSITE; PS50072; FORK_HEAD_36; 1.
 DR PROSITE; PS50073; FORK_HEAD_37; 1.
 DR PROSITE; PS50074; FORK_HEAD_38; 1.
 DR PROSITE; PS50075; FORK_HEAD_39; 1.
 DR PROSITE; PS50076; FORK_HEAD_40; 1.
 DR PROSITE; PS50077; FORK_HEAD_41; 1.
 DR PROSITE; PS50078; FORK_HEAD_42; 1.
 DR PROSITE; PS50079; FORK_HEAD_43; 1.
 DR PROSITE; PS50080; FORK_HEAD_44; 1.
 DR PROSITE; PS50081; FORK_HEAD_45; 1.
 DR PROSITE; PS50082; FORK_HEAD_46; 1.
 DR PROSITE; PS50083; FORK_HEAD_47; 1.
 DR PROSITE; PS50084; FORK_HEAD_48; 1.
 DR PROSITE; PS50085; FORK_HEAD_49; 1.
 DR PROSITE; PS50086; FORK_HEAD_50; 1.
 DR PROSITE; PS50087; FORK_HEAD_51; 1.
 DR PROSITE; PS50088; FORK_HEAD_52; 1.
 DR PROSITE; PS50089; FORK_HEAD_53; 1.
 DR PROSITE; PS50090; FORK_HEAD_54; 1.
 DR PROSITE; PS50091; FORK_HEAD_55; 1.
 DR PROSITE; PS50092; FORK_HEAD_56; 1.
 DR PROSITE; PS50093; FORK_HEAD_57; 1.
 DR PROSITE; PS50094; FORK_HEAD_58; 1.
 DR PROSITE; PS50095; FORK_HEAD_59; 1.
 DR PROSITE; PS50096; FORK_HEAD_60; 1.
 DR PROSITE; PS50097; FORK_HEAD_61; 1.
 DR PROSITE; PS50098; FORK_HEAD_62; 1.
 DR PROSITE; PS50099; FORK_HEAD_63; 1.
 DR PROSITE; PS50100; FORK_HEAD_64; 1.
 DR PROSITE; PS50101; FORK_HEAD_65; 1.
 DR PROSITE; PS50102; FORK_HEAD_66; 1.
 DR PROSITE; PS50103; FORK_HEAD_67; 1.
 DR PROSITE; PS50104; FORK_HEAD_68; 1.
 DR PROSITE; PS50105; FORK_HEAD_69; 1.
 DR PROSITE; PS50106; FORK_HEAD_70; 1.
 DR PROSITE; PS50107; FORK_HEAD_71; 1.
 DR PROSITE; PS50108; FORK_HEAD_72; 1.
 DR PROSITE; PS50109; FORK_HEAD_73; 1.
 DR PROSITE; PS50110; FORK_HEAD_74; 1.
 DR PROSITE; PS50111; FORK_HEAD_75; 1.
 DR PROSITE; PS50112; FORK_HEAD_76; 1.
 DR PROSITE; PS50113; FORK_HEAD_77; 1.
 DR PROSITE; PS50114; FORK_HEAD_78; 1.
 DR PROSITE; PS50115; FORK_HEAD_79; 1.
 DR PROSITE; PS50116; FORK_HEAD_80; 1.
 DR PROSITE; PS50117; FORK_HEAD_81; 1.
 DR PROSITE; PS50118; FORK_HEAD_82; 1.
 DR PROSITE; PS50119; FORK_HEAD_83; 1.
 DR PROSITE; PS50120; FORK_HEAD_84; 1.
 DR PROSITE; PS50121; FORK_HEAD_85; 1.
 DR PROSITE; PS50122; FORK_HEAD_86; 1.
 DR PROSITE; PS50123; FORK_HEAD_87; 1.
 DR PROSITE; PS50124; FORK_HEAD_88; 1.
 DR PROSITE; PS50125; FORK_HEAD_89; 1.
 DR PROSITE; PS50126; FORK_HEAD_90; 1.
 DR PROSITE; PS50127; FORK_HEAD_91; 1.
 DR PROSITE; PS50128; FORK_HEAD_92; 1.
 DR PROSITE; PS50129; FORK_HEAD_93; 1.
 DR PROSITE; PS50130; FORK_HEAD_94; 1.
 DR PROSITE; PS50131; FORK_HEAD_95; 1.
 DR PROSITE; PS50132; FORK_HEAD_96; 1.
 DR PROSITE; PS50133; FORK_HEAD_97; 1.
 DR PROSITE; PS50134; FORK_HEAD_98; 1.
 DR PROSITE; PS50135; FORK_HEAD_99; 1.
 DR PROSITE; PS50136; FORK_HEAD_100; 1.
 DR PROSITE; PS50137; FORK_HEAD_101; 1.
 DR PROSITE; PS50138; FORK_HEAD_102; 1.
 DR PROSITE; PS50139; FORK_HEAD_103; 1.
 DR PROSITE; PS50140; FORK_HEAD_104; 1.
 DR PROSITE; PS50141; FORK_HEAD_105; 1.
 DR PROSITE; PS50142; FORK_HEAD_106; 1.
 DR PROSITE; PS50143; FORK_HEAD_107; 1.
 DR PROSITE; PS50144; FORK_HEAD_108; 1.
 DR PROSITE; PS50145; FORK_HEAD_109; 1.
 DR PROSITE; PS50146; FORK_HEAD_110; 1.
 DR PROSITE; PS50147; FORK_HEAD_111; 1.
 DR PROSITE; PS50148; FORK_HEAD_112; 1.
 DR PROSITE; PS50149; FORK_HEAD_113; 1.
 DR PROSITE; PS50150; FORK_HEAD_114; 1.
 DR PROSITE; PS50151; FORK_HEAD_115; 1.
 DR PROSITE; PS50152; FORK_HEAD_116; 1.
 DR PROSITE; PS50153; FORK_HEAD_117; 1.
 DR PROSITE; PS50154; FORK_HEAD_118; 1.
 DR PROSITE; PS50155; FORK_HEAD_119; 1.
 DR PROSITE; PS50156; FORK_HEAD_120; 1.
 DR PROSITE; PS50157; FORK_HEAD_121; 1.
 DR PROSITE; PS50158; FORK_HEAD_122; 1.
 DR PROSITE; PS50159; FORK_HEAD_123; 1.
 DR PROSITE; PS50160; FORK_HEAD_124; 1.
 DR PROSITE; PS50161; FORK_HEAD_125; 1.
 DR PROSITE; PS50162; FORK_HEAD_126; 1.
 DR PROSITE; PS50163; FORK_HEAD_127; 1.
 DR PROSITE; PS50164; FORK_HEAD_128; 1.
 DR PROSITE; PS50165; FORK_HEAD_129; 1.
 DR PROSITE; PS50166; FORK_HEAD_130; 1.
 DR PROSITE; PS50167; FORK_HEAD_131; 1.
 DR PROSITE; PS50168; FORK_HEAD_132; 1.
 DR PROSITE; PS50169; FORK_HEAD_133; 1.
 DR PROSITE; PS50170; FORK_HEAD_134; 1.
 DR PROSITE; PS50171; FORK_HEAD_135; 1.
 DR PROSITE; PS50172; FORK_HEAD_136; 1.
 DR PROSITE; PS50173; FORK_HEAD_137; 1.
 DR PROSITE; PS50174; FORK_HEAD_138; 1.
 DR PROSITE; PS50175; FORK_HEAD_139; 1.
 DR PROSITE; PS50176; FORK_HEAD_140; 1.
 DR PROSITE; PS50177; FORK_HEAD_141; 1.
 DR PROSITE; PS50178; FORK_HEAD_142; 1.
 DR PROSITE; PS50179; FORK_HEAD_143; 1.
 DR PROSITE; PS50180; FORK_HEAD_144; 1.
 DR PROSITE; PS50181; FORK_HEAD_145; 1.
 DR PROSITE; PS50182; FORK_HEAD_146; 1.
 DR PROSITE; PS50183; FORK_HEAD_147; 1.
 DR PROSITE; PS50184; FORK_HEAD_148; 1.
 DR PROSITE; PS50185; FORK_HEAD_149; 1.
 DR PROSITE; PS50186; FORK_HEAD_150; 1.
 DR PROSITE; PS50187; FORK_HEAD_151; 1.
 DR PROSITE; PS50188; FORK_HEAD_152; 1.
 DR PROSITE; PS50189; FORK_HEAD_153; 1.
 DR PROSITE; PS50190; FORK_HEAD_154; 1.
 DR PROSITE; PS50191; FORK_HEAD_155; 1.
 DR PROSITE; PS50192; FORK_HEAD_156; 1.
 DR PROSITE; PS50193; FORK_HEAD_157; 1.
 DR PROSITE; PS50194; FORK_HEAD_158; 1.
 DR PROSITE; PS50195; FORK_HEAD_159; 1.
 DR PROSITE; PS50196; FORK_HEAD_160; 1.
 DR PROSITE; PS50197; FORK_HEAD_161; 1.
 DR PROSITE; PS50198; FORK_HEAD_162; 1.
 DR PROSITE; PS50199; FORK_HEAD_163; 1.
 DR PROSITE; PS50200; FORK_HEAD_164; 1.
 DR PROSITE; PS50201; FORK_HEAD_165; 1.
 DR PROSITE; PS50202; FORK_HEAD_166; 1.
 DR PROSITE; PS50203; FORK_HEAD_167; 1.
 DR PROSITE; PS50204; FORK_HEAD_168; 1.
 DR PROSITE; PS50205; FORK_HEAD_169; 1.
 DR PROSITE; PS50206; FORK_HEAD_170; 1.
 DR PROSITE; PS50207; FORK_HEAD_171; 1.
 DR PROSITE; PS50208; FORK_HEAD_172; 1.
 DR PROSITE; PS50209; FORK_HEAD_173; 1.
 DR PROSITE; PS50210; FORK_HEAD_174; 1.
 DR PROSITE; PS50211; FORK_HEAD_175; 1.
 DR PROSITE; PS50212; FORK_HEAD_176; 1.
 DR PROSITE; PS50213; FORK_HEAD_177; 1.
 DR PROSITE; PS50214; FORK_HEAD_178; 1.
 DR PROSITE; PS50215; FORK_HEAD_179; 1.
 DR PROSITE; PS50216; FORK_HEAD_180; 1.
 DR PROSITE; PS50217; FORK_HEAD_181; 1.
 DR PROSITE; PS50218; FORK_HEAD_182; 1.
 DR PROSITE; PS50219; FORK_HEAD_183; 1.
 DR PROSITE; PS50220; FORK_HEAD_184; 1.
 DR PROSITE; PS50221; FORK_HEAD_185; 1.
 DR PROSITE; PS50222; FORK_HEAD_186; 1.
 DR PROSITE; PS50223; FORK_HEAD_187; 1.
 DR PROSITE; PS50224; FORK_HEAD_188; 1.
 DR PROSITE; PS50225; FORK_HEAD_189; 1.
 DR PROSITE; PS50226; FORK_HEAD_190; 1.
 DR PROSITE; PS50227; FORK_HEAD_191; 1.
 DR PROSITE; PS50228; FORK_HEAD_192; 1.
 DR PROSITE; PS50229; FORK_HEAD_193; 1.
 DR PROSITE; PS50230; FORK_HEAD_194; 1.
 DR PROSITE; PS50231; FORK_HEAD_195; 1.
 DR PROSITE; PS50232; FORK_HEAD_196; 1.
 DR PROSITE; PS50233; FORK_HEAD_197; 1.
 DR PROSITE; PS50234; FORK_HEAD_198; 1.
 DR PROSITE; PS50235; FORK_HEAD_199; 1.
 DR PROSITE; PS50236; FORK_HEAD_200; 1.
 DR PROSITE; PS50237; FORK_HEAD_201; 1.
 DR PROSITE; PS50238; FORK_HEAD_202; 1.
 DR PROSITE; PS50239; FORK_HEAD_203; 1.
 DR PROSITE; PS50240; FORK_HEAD_204; 1.
 DR PROSITE; PS50241; FORK_HEAD_205; 1.
 DR PROSITE; PS50242; FORK_HEAD_206; 1.
 DR PROSITE; PS50243; FORK_HEAD_207; 1.
 DR PROSITE; PS50244; FORK_HEAD_208; 1.
 DR PROSITE; PS50245; FORK_HEAD_209; 1.
 DR PROSITE; PS50246; FORK_HEAD_210; 1.
 DR PROSITE; PS50247; FORK_HEAD_211; 1.
 DR PROSITE; PS50248; FORK_HEAD_212; 1.
 DR PROSITE; PS50249; FORK_HEAD_213; 1.
 DR PROSITE; PS50250; FORK_HEAD_214; 1.
 DR PROSITE; PS50251; FORK_HEAD_215; 1.
 DR PROSITE; PS50252; FORK_HEAD_216; 1.
 DR PROSITE; PS50253; FORK_HEAD_217; 1.
 DR PROSITE; PS50254; FORK_HEAD_218; 1.
 DR PROSITE; PS50255; FORK_HEAD_219; 1.
 DR PROSITE; PS50256; FORK_HEAD_220; 1.
 DR PROSITE; PS50257; FORK_HEAD_221; 1.
 DR PROSITE; PS50258; FORK_HEAD_222; 1.
 DR PROSITE; PS50259; FORK_HEAD_223; 1.
 DR PROSITE; PS50260; FORK_HEAD_224; 1.
 DR PROSITE; PS50261; FORK_HEAD_225; 1.
 DR PROSITE; PS50262; FORK_HEAD_226; 1.
 DR PROSITE; PS50263; FORK_HEAD_227; 1.
 DR PROSITE; PS50264; FORK_HEAD_228; 1.
 DR PROSITE; PS50265; FORK_HEAD_229; 1.
 DR PROSITE; PS50266; FORK_HEAD_230; 1.
 DR PROSITE; PS50267; FORK_HEAD_231; 1.
 DR PROSITE; PS50268; FORK_HEAD_232; 1.
 DR PROSITE; PS50269; FORK_HEAD_233; 1.
 DR PROSITE; PS50270; FORK_HEAD_234; 1.
 DR PROSITE; PS50271; FORK_HEAD_235; 1.
 DR PROSITE; PS50272; FORK_HEAD_236; 1.
 DR PROSITE; PS50273; FORK_HEAD_237; 1.
 DR PROSITE; PS50274; FORK_HEAD_238; 1.
 DR PROSITE; PS50275; FORK_HEAD_239; 1.
 DR PROSITE; PS50276; FORK_HEAD_240; 1.
 DR PROSITE; PS50277; FORK_HEAD_241; 1.
 DR PROSITE; PS50278; FORK_HEAD_242; 1.
 DR PROSITE; PS50279; FORK_HEAD_243; 1.
 DR PROSITE; PS50280; FORK_HEAD_244; 1.
 DR PROSITE; PS50281; FORK_HEAD_245; 1.
 DR PROSITE; PS50282; FORK_HEAD_246; 1.
 DR PROSITE; PS50283; FORK_HEAD_247; 1.
 DR PROSITE; PS50284; FORK_HEAD_248; 1.
 DR PROSITE; PS50285; FORK_HEAD_249; 1.
 DR PROSITE; PS50286; FORK_HEAD_250; 1.
 DR PROSITE; PS50287; FORK_HEAD_251; 1.
 DR PROSITE; PS50288; FORK_HEAD_252; 1.
 DR PROSITE; PS50289; FORK_HEAD_253; 1.
 DR PROSITE; PS50290; FORK_HEAD_254; 1.
 DR PROSITE; PS50291; FORK_HEAD_255; 1.
 DR PROSITE; PS50292; FORK_HEAD_256; 1.
 DR PROSITE; PS50293; FORK_HEAD_257; 1.
 DR PROSITE; PS50294; FORK_HEAD_258; 1.
 DR PROSITE; PS50295; FORK_HEAD_259; 1.
 DR PROSITE; PS50296; FORK_HEAD_260; 1.
 DR PROSITE; PS50297; FORK_HEAD_261; 1.
 DR PROSITE; PS50298; FORK_HEAD_262; 1.
 DR PROSITE; PS50299; FORK_HEAD_263; 1.
 DR PROSITE; PS50300; FORK_HEAD_264; 1.
 DR PROSITE; PS50301; FORK_HEAD_265; 1.
 DR PROSITE; PS50302; FORK_HEAD_266; 1.
 DR PROSITE; PS50303; FORK_HEAD_267; 1.
 DR PROSITE; PS50304; FORK_HEAD_268; 1.
 DR PROSITE; PS50305; FORK_HEAD_269; 1.
 DR PROSITE; PS50306; FORK_HEAD_270; 1.
 DR PROSITE; PS50307; FORK_HEAD_271; 1.
 DR PROSITE; PS50308; FORK_HEAD_272; 1.
 DR PROSITE; PS50309; FORK_HEAD_273; 1.
 DR PROSITE; PS50310; FORK_HEAD_274; 1.
 DR PROSITE; PS50311; FORK_HEAD_275; 1.
 DR PROSITE; PS50312; FORK_HEAD_276; 1.
 DR PROSITE; PS50313; FORK_HEAD_277; 1.
 DR PROSITE; PS50314; FORK_HEAD_278; 1.
 DR PROSITE; PS50315; FORK_HEAD_279; 1.
 DR PROSITE; PS50316; FORK_HEAD_280; 1.
 DR PROSITE; PS50317; FORK_HEAD_281; 1.
 DR PROSITE; PS50318; FORK_HEAD_282; 1.
 DR PROSITE; PS50319; FORK_HEAD_283; 1.
 DR PROSITE; PS50320; FORK_HEAD_284; 1.
 DR PROSITE; PS50321; FORK_HEAD_285; 1.
 DR PROSITE; PS50322; FORK_HEAD_286; 1.
 DR PROSITE; PS50323; FORK_HEAD_287; 1.
 DR PROSITE; PS50324; FORK_HEAD_288; 1.
 DR PROSITE; PS50325; FORK_HEAD_289; 1.
 DR PROSITE; PS50326; FORK_HEAD_290; 1.
 DR PROSITE; PS50327; FORK_HEAD_291; 1.
 DR PROSITE; PS50328; FORK_HEAD_292; 1.
 DR PROSITE; PS50329; FORK_HEAD_293; 1.
 DR PROSITE; PS50330; FORK_HEAD_294; 1.
 DR PROSITE; PS50331; FORK_HEAD_295; 1.
 DR PROSITE; PS50332; FORK_HEAD_296; 1.
 DR PROSITE; PS50333; FORK_HEAD_297; 1.
 DR PROSITE; PS50334; FORK_HEAD_298; 1.
 DR PROSITE; PS50335; FORK_HEAD_299; 1.
 DR PROSITE; PS50336; FORK_HEAD_300; 1.
 DR PROSITE; PS50337; FORK_HEAD_301; 1.
 DR PROSITE; PS50338; FORK_HEAD_302; 1.
 DR PROSITE; PS50339; FORK_HEAD_303; 1.
 DR PROSITE; PS50340; FORK_HEAD_304; 1.
 DR PROSITE; PS50341; FORK_HEAD_305; 1.
 DR PROSITE; PS50342; FORK_HEAD_306; 1.
 DR PROSITE; PS50343; FORK_HEAD_307; 1.
 DR PROSITE; PS50344; FORK_HEAD_308; 1.
 DR PROSITE; PS50345; FORK_HEAD_309; 1.
 DR PROSITE; PS50346; FORK_HEAD_310; 1.
 DR PROSITE; PS50347; FORK_HEAD_311; 1.
 DR PROSITE; PS50348; FORK_HEAD_312; 1.
 DR PROSITE; PS50349; FORK_HEAD_313; 1.
 DR PROSITE; PS50350; FORK_HEAD_314; 1.
 DR PROSITE; PS50351; FORK_HEAD_315; 1.
 DR PROSITE; PS50352; FORK_HEAD_316; 1.
 DR PROSITE; PS50353; FORK_HEAD_317; 1.
 DR PROSITE; PS50354; FORK_HEAD_318; 1.
 DR PROSITE; PS50355; FORK_HEAD_319; 1.
 DR PROSITE; PS50356; FORK_HEAD_320; 1.
 DR PROSITE; PS50357; FORK_HEAD_321; 1.
 DR PROSITE; PS50358; FORK_HEAD_322; 1.
 DR PROSITE; PS50359; FORK_HEAD_323; 1.
 DR PROSITE; PS50360; FORK_HEAD_324; 1.
 DR PROSITE; PS50361; FORK_HEAD_325; 1.
 DR PROSITE; PS50362; FORK_HEAD_326; 1.
 DR PROSITE; PS50363; FORK_HEAD_327; 1.
 DR PROSITE; PS50364; FORK_HEAD_328; 1.
 DR PROSITE; PS50365; FORK_HEAD_329;

MEDLINE=92265309; PubMed=1350202;

RA Tao W.; Lai E.;

RT "Telencephalon-restricted expression of BF-1, a new member of the HNF-3/Fork head gene family, in the developing rat brain.";

RL Neuron 8: 957-966(1992).

CC -!- FUNCTION: Plays an important role in the establishment of the regional subdivision of the developing brain and in the development of the telencephalon. Sequence specific DNA-binding protein with a distinct binding specificity.

CC -!- SUBCELLULAR LOCATION: Nuclear.

CC -!- TISSUE SPECIFICITY: Is expressed in the cortex, olfactory bulb, hippocampus, and caudate putamen.

CC -!- DEVELOPMENTAL STAGE: More abundant in fetal brain compared with the adult brain.

CC -!- SIMILARITY: Contains 1 fork-head domain.

CC ---

This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to license@isb-sib.ch).

CC ---

DR EMBL; MB7634; AAA0812.1; -.

DR PIR; JH0672; JH0672.

DR HSSP; Q6245; 2HFF.

DR TRANSFAC; T01831; -.

DR InterPro; IPR001766; TF_Fork_head.

DR PF00250; Fork_head; 1.

DR PRINTS; PR00053; FORKHEAD.

DR Prodrom; PD000425; TF_Fork_head; 1.

DR SMART; SR00339; FH_1.

DR PROSITE; PS000657; FORK_HEAD_1.

DR PROSITE; PS00058; FORK_HEAD_2; 1.

DR PROSITE; PS50039; FORK_HEAD_3; 1.

KW Transcription regulation; DNA binding; Nuclear protein; Developmental protein.

FT DOMAIN 33 56 HIS-RICH.

FT DOMAIN 57 82 PRO-RICH.

FT DOMAIN 70 74 POLY-GLN.

FT DNA BIND 171 262 FORK_HEAD.

SEQUENCE 480 AA; 51466 MW; 00425B2F5298F444 CRC64;

Query Match Score 188.5; DB 1; Length 480;

Best Local Similarity 43.9%; Pred. No. 2.5e-13;

Matches 36; Conservative 20; Mismatches 19; Indels 7; Gaps 2;

DR SMART; SM00339; FH; 1.

DR EMBL; BC046958; AAH46958; 2; -.

DR HSSP; Q63245; 2HFF.

DR TRANSFAC; T02353; -.

DR MGII; MGII_134746; Foxg1.

DR GO; GO:0007420; P:brain development; IMP.

DR GO; GO:0030182; P:neuron differentiation; IMP.

DR InterPro; IPR001766; TF_Fork_head.

DR Pfam; PF00250; Fork_head; 1.

DR PRINTS; PR00053; FORKHEAD.

DR ProDom; PD000425; TF_Fork_head; 1.

DR SMART; SM00339; FH; 1.

DR PROSITE; PS000657; FORK_HEAD_1.

DR PROSITE; PS000658; FORK_HEAD_2; 1.

DR PROSITE; PS50039; FORK_HEAD_3; 1.

KW Transcription regulation; DNA-binding; Nuclear protein;

FT DOMAIN 33 56 HIS-RICH.

FT DOMAIN 57 83 PRO-RICH.

FT DOMAIN 70 75 POLY-GLN.

FT DNA BIND 172 263 FORK_HEAD.

SQ SEQUENCE 481 AA; 51624 MW; FF6786C3513AB452 CRC64;

RESULT 15

FXGB_MOUSE STANDARD PRT; 481 AA.

ID FXGB_MOUSE STANDARD PRT; 481 AA.

AC Q60987; 35, Created)

DT 01-NOV-1997 (Rel. 35, Last sequence update)

DT 15-MAR-2004 (Rel. 43, Last annotation update)

DE Forkhead box protein G1B (Forkhead-related protein FKHL1).

DE (Transcription factor BF-1) (Brain factor 1) (BF1).

GN FOXG1B OR FOXG1 OR FKHL1 OR HFFBF1.

OS Mus musculus (Mouse); Chordata; Craniata; Vertebrata; Butteleostomi;

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Butteleostomi;

OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.

OX NCBI_TaxID=10900;

[1] _SEQUENCE FROM N.A.

RP

QY 14 SYAELITTA MASPEKRLTIAQVYEWMMVQNVPYFRDKGDSNSAGWKNSIRHNLSTLHSRF 73
Db 177 SYNALIMMAIATROSPEKLTLINGLIEFIMKNFPLYRE---NKQGWONSIRHNLSTLNKCF 231
QY 74 MRIQN-EGAGKSSWVINPDA 93
Db :::: : VKVPRHYDPFGKGNYWMLDSS 253

Search completed: July 23, 2004, 10:24:37
Job time : 14 secs

This page blank (uspto)